

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:26:31 ; Search time 6206.58 Seconds  
(without alignments)  
8107.924 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135

Perfect score: 4683  
Sequence: 1 atgtgactactgctctctc.....gtctgcgaggaagcagcc 4683

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estcom:\*  
5: em\_estcpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qd\_est1:\*  
11: qd\_est2:\*  
12: qd\_hic:\*  
13: qd\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vct:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382.4	12.4	653	11	BG819678 602783172
2	405.8	8.7	499	11	B1185857
3	348.8	7.4	455	11	BF544715
4	343.8	7.3	355	10	AJ003472
5	341.2	7.3	458	10	A1454704
6	282	6.0	487	10	BG187880
7	274.2	5.9	410	10	AM529324
8	271.2	5.8	427	10	AM159035
9	254.8	5.4	483	11	BG307959
10	254	5.4	744	11	BF348620
11	244	5.2	344	13	A2576100
12	234.8	5.0	1232	12	AK015320

13	234.8	5.0	1376	12	AK014832
14	228.4	4.9	468	11	BF565509
15	219.8	4.7	279	10	BB336040
16	217.6	4.6	291	10	BB333216
17	216.2	4.6	355	10	BB461534
18	202.4	4.3	252	10	BB381728
19	202.4	4.3	287	11	BF928859
20	202.2	4.3	270	10	BB336019
21	201	4.1	248	10	BB335137
22	191.2	4.1	296	10	BB329941
23	188.4	4.0	247	10	BB330789
24	188	4.0	280	10	BB331659
25	182.8	3.9	257	10	BB330426
26	179.6	3.8	610	13	AZ56759
27	179.4	3.8	253	10	BB331552
28	178.6	3.8	304	10	BR280255
29	175.6	3.7	619	13	FR0002941
30	172.6	3.7	286	10	BB333970
31	169.8	3.6	272	10	BB330770
32	169	3.6	252	10	BB334670
33	167.8	3.6	239	10	BB332000
34	163.2	3.5	402	10	AM048129
35	154	3.3	816	13	CNS031Y7
36	150.6	3.2	849	13	CNS031Y6
37	149.2	3.2	3133	12	AK016518
38	148.6	3.2	216	10	BB328887
39	147.4	3.1	409	10	AM014717
40	144.6	3.1	219	10	BB342624
41	143	3.1	147	13	HSKC18B12
42	140.8	3.0	912	13	CNS021TG
43	138	2.9	491	13	AZ657601
44	136.2	2.9	212	10	BB330864
45	135.4	2.9	967	13	CNS025E3

#### ALIGNMENTS

RESULT 1  
LOCUS BG819678 653 bp mRNA  
DEFINITION 602783172P1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4934248  
5' mRNA sequence.  
ACCESSION BG819678  
KEYWORDS  
SOURCE  
ORGANISM Human  
Tissue: Fetus  
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
REFERENCE  
AUTHORS NIH-MCC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-rt@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM10862 row: 0 column: 17  
High quality sequence stop: 648.  
Location/Qualifiers  
1. 653  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4934248"  
/clone\_lib="NCI\_CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"



Db	423	AAATTCCTCCAGCCATGACCAATTCACCAAGAGCATTCATATTCCTGCTCCACCCCTTT	482
Qy	3338	ccaaagaaacgttgaat 3354 	
Db	483	CCAGAGAGCCTTGAAAT 499	
RESULT	3		
LOCUS	BF544715		
DEFINITION	BF544715 455 bp mRNA EST 11-DEC-2000		
ACCESSION	U1-R-BT0-qk-h-03-0-01.1 U1-R-BT0 Rattus norvegicus cDNA clone		
VERSION	U1-R-BT0-qk-h-03-0-01.5', mRNA sequence.		
KEYWORDS	BF544715 BF544715.1 GI:11635822		
SOURCE	EST.		
ORGANISM	Norway rat. Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 455)		
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE ID- 1788947 Seq primer: M3 forward. location/Qualifiers 1..455 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="U1-R-BT0-qk-h-03-0-01" /clone_11b="U1-R-BT0" /dev_stage="adult" /lab_host="DH10B (Life Technologies)" /note="Vector: pTZ19D-lac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; This library (U1-R-BT0) consists of a mixture of individually lagged normalized libraries constructed from rat hippocampus, thalamus, mid-brain, medulla, corpus striatum, cerebral cortex and testis. The lag used to identify the source tissue is a string of 3-6 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. This library was then subtracted using a driver consisting of a mixture of all clones from U1-R-A0, U1-R-A1, U1-R-E0, U1-R-E1, U1-R-C0, U1-R-C1, U1-R-C2 and U1-R-C2P."		
BASE COUNT	108 a 134 c 109 g 104 t		
ORIGIN			
Query Match	7.4%; Score 348.8; DB 11; Length 455;		
Best local Similarity	86.2%; Pred. No. 2.6e-87;		
Matches 386; Conservative 0; Mismatches 62; Indels 0; Caps 0			
Qy	1568	ltaacgtgtatgtatgcatacgtatlaactcataaatgtatataactctaac 1627 	
Db	8	TCACATGCGCTGTGATGTGGCTACCTTATTTATGCGCATGCAAGATGCTTAAAT 67	
Qy	1628	tacttctttaaaccacgcgcgaatgtgacatlltgaaatcagatgccttlaaacctlltgcg 1687	

Db	68	TGCTTCCTTTCAACCAACCCGAGCTGGCGTTTGAGAACATGGGACTCTGACACTCTACG	127
Oy	1688	ctgtgcacaaagaagaatggacgaaggagatcacacgtgcacagtglttggttcaaccacaac	1747
Db	128	ATGTCAGCAAGCAAGCTTGACGAGGAGACTACACATGTAATGTGGTGGGCGACCAACG	187
Oy	1748	ctctcaccagccgcgcgcgtctccacgtgcagcttgaaagcttcgcgccttcatacaaccctctg	1807
Db	188	TCTTCACCAACCCAGAGAGTTCACGTACGTCAAGTCAAACTTCTCTCTTTATCCAAACCTTTG	247
Oy	1808	agttctccaaattctccattggacacacgagttctcatctccctgtgtgtgctcagaagg	1867
Db	248	AGTTCGACAGACTTCTCTATTCGCTGACCGGCGTCTTCATCCCATGTGTGGTGCTCGGGG	307
Oy	1868	acttaccatcacgaltcaacctggcagaagaatggccgcgaatccctcggagaccttgagg	1927
Db	308	ACTTACCTCATTCACCATCACTCCGACAGAGATGCGCGGCGATTCACGCAACCTCTGGTG	367
Oy	1928	tgaccatctgacaataltgaaactcaagagctcctctggagattcccaatcctctgcgtcatgc	1987
Db	368	TAAACCATTCATACATGACCTTCACTCACTCTCTGAGAGATCTCCAACTCTCACTGATCC	427
Oy	1988	acaaatggaatatacacttgcataagccg	2015
Db	428	ACAAGCGCAATTATACCTTGTATGCGCGG	455

[illegible]







Fax: 516 367 8874  
Email: mcomble@cs.hl.org  
Plate: zaf50 row: e column: 05  
Seq primer: M13 universal forward primer  
High quality sequence stop: 427.  
Location/Qualifiers

## FEATURES

1. 427  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone\_lib="zaf50e05"  
/clone\_lib="Xenopus EST library"  
/rname="total brain tissue"  
/cell\_line="W22-TGA"  
/dev\_stage="tadpole"  
/note="Vector: Lambda zap I; Site\_1: XbaI; This library was supplied by Holly Cline (Cold Spring Harbor Labs). cDNA synthesis with oligo dt Xba I (Xba I cloning site). RNA: stage 50-56 tadpoles, total brain tissue, CTC extraction method."

3 COUNT 114 a 122 c 91 g 100 t

Query Match 5.8%; Score 271.2; DB 10; Length 427;  
Best Local Similarity 79.8%; Pred. No. 2.3e-65;  
Matches 335; Conservative 0; Mismatches 78; Indels 7; Gaps 1;

1226 ttttttcgttcttgatgaagaagtgatgagtcacagagagccggttcccttattgca 1285

15 tcatctctgcatTTAGGAGAGAGTGTGAGCCCGAGAACCTGTCCTTATGTGCA 74

1286 acgtgaaggaacacgttctgcccagatcagtgagcccttgagcgttcccttcc 1345

75 ACCTAAAGAGAGACCCCTTCCGACATCTTGACGACCTGAGCGACGAGCCCATTTATCA 134

1346 aggggtgagcagtcacagcagcagcagcagcagcagcagcagcagcagcagc 1405

135 AGAAGCGCCACACAGATAGTACGTTATACCTCTGAGGCGACGTTGACGTTACG 194

1406 tgaacatctcagcgtccaggtcagcagcagcagcagcagcagcagcagcagc 1465

195 TGACATCACCACACATCAAGTCCGAGATGCGGGGCTGACCGATGCGGCTGAC 254

1466 cagtgagagatctgctctgtagcagcagcagcagcagcagcagcagcagcagc 1525

255 CTGCGCGGCTGCTCTTGTACCGAGCTCGAATTAACCT-----CTGCGAAGTATGCTGC 307

1526 caatgaacacatcacagcaatagcagagcagcagcagcagcagcagcagcagc 1585

308 CAATGAACAAACATCAGCGCCATACGCTGCGGCGACATCATCCAGTCCGCTCATTTG 367

1586 gctatcgtatctcacttaataatggtacagcaaaccttaacgtctctccttcaaccac 1645

368 GATATCCCTACTACTCATCAAGTGTGTAACAAACCTTAACCTCTCTCCATTTATATCACC 427

RESULT 9

LOCUS BG307959 483 bp mRNA EST 22-FEB-2001

DEFINITION fms5b04.y1 zebrafish adult retina cDNA Danio rerio cDNA c. ne

MOLECULE : mRNA sequence.

ACCESSION BG307959 GI:13105486

VERSION BG307959.1 GI:13105486

KEYWORDS EST.

SOURCE zebrafish.

ORGANISM Danio rerio

REFERENCE 1 (bases 1 to 483)

AUTHORS Clark, M., Johnson, S.L., Lehrbach, H., Lee, K., Li, F., Morris, M., Ruddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

R., Steptoe, M., Theislow, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Rutter, E., Kohn, S., Shin, T., Jackson, T., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Washu zebrafish EST Project 1998  
Unpublished (1998)  
Contact: Stephen L. Johnson  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbratfish@wustl.edu

## FEATURES

1. 483  
/organism="Danio rerio"  
/strain="wild-type"  
/db\_xref="taxon:7955"  
/clone="419358"  
/clone\_lib="Zebrafish adult retina cDNA"  
/sex="mixed"  
/dev\_stage="1-2 years"  
/lab\_host="E.Coli XL1-Blue MRF" (XL1-Blue MRF)  
/note="Vector: Lambda ZAP II (pBluescript SK-); Site\_1: EcoRI; Site\_2: SalI; This zebrafish library was constructed by Dr. Susan E. Brockerhoff (email: sbrocker@u.washington.edu) RZPD library number: 760"

## BASE COUNT

112 a 130 c 114 g 107 t

Query Match 5.4%; Score 254.8; DB 11; Length 483;  
Best Local Similarity 73.2%; Pred. No. 1.1e-60;  
Matches 353; Conservative 0; Mismatches 127; Indels 2; Gaps 2;

3891 tccatgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 3950

1 TCCTCGATGAGAGGACATGCTGCTTCATGCAAGACAGTCGGAGACCCCTCAACAT 60

3951 ccaatgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4010

61 CAATGAGATGAGAGGACATGCTGCTTCATGCAAGACAGTCGGAGACCCCTCAACAT 120

4011 gctcttctgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4070

121 CGCTCACAAACAAATGACGCTTGTATATGCTGCAACAGACAGATTCGGGTATTA 180

4071 gactgcatgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4130

181 CAGATGCTGCTGCAAGCAACATGCGGCTTCGATGAGATCATGCTTAATCTTACGTCA 240

4131 agtctgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4188

241 AGTACACACGAGCAACGACGCTGCTGCAACAGACAGATTCGGGTATTA 300

4189 cttcttgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4248

301 CTCTTCTGAGTCCCGGACCAATGCGGAGCTCCATGACGAGCGATTTTACAGTAC 360

4249 tccgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4308

361 TCAGAGCATTAACACTGAAGATGCGGAGTTCATGACGACCCGAGTCCCTTTAC 420

4309 cgtcttgagatgaagaacatgctctgctgtagagcgtgttgaggacccctctcagcgt 4368

421 CGTTCGAGATTAACACTGAAGATGCGGAGTTCATGACGACCCGAGTCCCTTTAC 480

4369 gtc 4370





Db	770	GTGCTTCATCCCTCTTCAAGTACGACGAGATACGTTACGTCGTCGTTCTTGAGAGAGACAC	829
Db	492	gcttc 497	
Db	830	GGTCTC 835	
RESULT	13		
LOCUS	AK014832		
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921507G06, full insert sequence.		
ACCESSION	AK014832		
VERSION	AK014832.1	GI:12852903	
KEYWORDS	CAP TRAPPER.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:11b:RIKEN full-length enriched mouse cDNA library clone:4921507G06.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus		
AUTHORS	1 (bases 1 to 1376)		
TITLE	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Methods in enzymology. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2 (bases 1 to 1376)		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtriction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome research. 10 (10), 1617-1630 (2000)		
PUBMED	20499374		
REFERENCE	11042159		
AUTHORS	3 (bases 1 to 1376)		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
JOURNAL	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multichipillary sequencer		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	20530913		
REFERENCE	11076861		
AUTHORS	4 (bases 1 to 1376)		
TITLE	The RIKEN genome exploration research group phase II team and the FANTOM Consortium.		
JOURNAL	Functional annotation of a full-length mouse cDNA collection		
MEDLINE	Nature 409, 685-690 (2001)		
PUBMED	5 (bases 1 to 1376)		
REFERENCE	11076861		
AUTHORS	Aachisi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arawaka, T., Harai, A., Hayatsu, N., Hi,amoto, K., Hirakawa, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurikawa, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Saito-cho, Tsukuba, Ibaraki, 305-3858, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/). Tel: 81-45-503-9222.		

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:29:31 ; Search time 578.64 Seconds  
(without alignments)  
6938.432 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135  
Perfect score: 4683  
Sequence: 1 atgtgatactgtctctc.....gtctgcgcgaagaagcagc 4683

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101:.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT:.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT:.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT:.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT:.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT:.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT:.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT:.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT:.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT:.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT:.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT:.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT:.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT:.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT:.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT:.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT:.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT:.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT:.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT:.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT:.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT:.\*  
23: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4683	100.0	6604	19	AAV31981 Human Down syndrom
2	4678.8	99.9	6413	19	AAV31988 Human Down syndrom
3	1806.8	38.6	5824	22	AAI58196 Human polynucleoti
4	636	13.6	842	19	AAV31985 Mouse Down syndrom
5	529.2	11.3	1493	19	AAV27207 cDNA clone ethb001
6	502	10.7	898	19	AAV31986 Mouse Down syndrom
7	299.4	6.4	2173	19	AAV31987 Human polynucleoti
8	227.6	4.9	471	22	AAI59982 Human polynucleoti
9	139	3.0	388	19	AAV31982 Human Down syndrom
10	91.4	2.0	3617	22	AAH24799 Nucleotide sequenc
11	91.4	2.0	4336	22	AAH05447 Human secreted pro

12	78	1.7	808	22	AAH06899
13	78	1.7	2537	22	AAH14602
14	66.4	1.4	6000	16	AAO86478 Human PTP-OR. Hom
15	66.4	1.4	6000	18	AAH85389 Human protein tyro
16	66.4	1.4	6000	20	AAH06095 Human protein tyro
17	66.4	1.4	6000	22	AAH02144 Rat hypothetical 1
18	66.2	1.4	2401	22	AAH0107 Human protein tyro
19	57	1.2	7702	21	AAH8739 Human cDNA encodin
20	57	1.2	7702	21	AAZ59133 Human cDNA encodin
21	57	1.2	7705	22	AAH98405 Human EST-derived
22	57	1.2	7705	22	AAH98405 Human cDNA encodin
23	57	1.2	7705	22	AAH98405 Human cDNA encodin
24	57	1.2	7741	22	AAH22448 Drosophila sp. ROB
25	53.2	1.1	4188	20	AAH5767 Rat oploid recepto
26	53.2	1.1	4188	20	AAH5767 Rat oploid recepto
27	49.6	1.1	2179	14	AAH051017 Rat oploid recepto
28	49.6	1.1	2337	14	AAH051016 Rat oploid recepto
29	49.6	1.1	306	14	AAH051015 Mechanical stress
30	48	1.0	837	21	AAH26301 Mechanical stress
31	48	1.0	7679	21	AAH26322 Mechanical stress
32	48	1.0	8882	21	AAH26321 Mechanical stress
33	47.8	1.0	3774	22	AAH01598 Human LI cell adhe
34	46.4	1.0	7710	22	AAH44662 Rat receptor type-
35	45.8	1.0	5690	16	AAH086902 Human UNC-40 nucle
36	45.2	1.0	4344	22	AAH090961 Deleted in colorec
37	45.2	1.0	4608	16	AAH080196 Human folliculin-
38	45.2	1.0	4608	20	AAH23431 Human polynucleoti
39	45.2	1.0	5186	12	AAH012752 Human polynucleoti
40	45.2	1.0	2526	22	AAH30917 Human polynucleoti
41	44.6	1.0	3730	22	AAH159396 Human polynucleoti
42	44.6	1.0	3783	22	AAH161182 Human polynucleoti
43	44.6	1.0	4843	22	AAH37553 Human polynucleoti
44	44.6	1.0	5202	22	AAH99637 Human protein encod
45	44.2	0.9	5994	15	AAH05674 Sequence encoding

## ALIGNMENTS

AAV31981	RESULT 1
AAV31981 standard; cDNA: 6604 bp.	
AAV31981:	
28-SEP-1998 (first entry)	
Human Down syndrome--cell adhesion molecule DS-CAM) cDNA.	
US-CAM1: Down syndrome--cell adhesion molecule; neural cell;	
signal transduction; trisomy 21; mental retardation;	
holoprosencephaly; corpus callosum agenesis;	
schizencephaly; diagnosis: assay: human; ds: ss.	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT CDS	453..6185
FT FT	/*tag- a
FT sig-peptide	452..521
FT FT	/*tag- b
FT mat-peptide	522..6162
FT FT	/*tag- c
XX	
XX	
PN W09817795-A1.	
XX	
PD 30-APR-1998.	
XX	
PF 23-OCT-1997; 97WO-US19547.	
XX	
PR 25-OCT-1996; 96US-0029322.	
XX	
PA (CEDA-) CINDARS SINAI MEDICAL CENT.	



XX Korenberg JR:  
 XX  
 XX WPI: 1998-271791/24.  
 UR  
 DR P-PSDB: AAM42086.  
 XX  
 XX New Isolated Down's Syndrome-cell adhesion molecule - used to  
 PT develop products for detection, diagnosis and therapy of  
 PT developmental and neurological abnormalities  
 XX  
 XX Claim 3; Page 65-72; 109pp; English.

This cDNA clone codes for Down syndrome-cell adhesion molecule-DS-CAM1 (see AAMW42086), a cell surface glycoprotein belonging to a novel subclass of the Ig superfamily with highest homology to neural cell adhesion molecules. A modified direct cDNA selection technique was applied to bacterial and P1 artificial chromosomes between ET52 and MX1 by using cDNA from trisomy 21 human foetal brain. A unique cDNA fragment, designated E51 (see AAV11982), was obtained and used to screen a trisomy 21 human foetal brain (14-wk gestation) cDNA library. Further clones were obtained by exon trapping. The complete DS-CAM1 cDNA sequence was deduced from overlapping clones. A splice variant cDNA (see AAV13988), encoding non-membrane bound DS-CAM2 (see AAMW42087), was also identified. The DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome 21q22.2-22.3. The invention also provides murine DS-CAM partial transcripts (see AAV13985-87), expression vectors and host cells, transgenic animals, antisense oligonucleotides, and primers useful for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are associated with developmental and neurological processes. They can be used in e.g. neural prosthetic devices used in entubulation methods of repairing (regenerating) damaged or severed peripheral nerves. The products can also be used in detection, diagnosis and therapy of developmental and neurological abnormalities such as Down syndrome, mental retardation, holoprosencephaly, agenesis of the corpus callosum, or schizencephaly. Antisense oligonucleotides are used for inhibiting translation of mRNA.

Sequence 6604 BP; 1673 A; 1827 C; 1768 G; 1336 T; 0 other;

Query Match	Similarity	Score	4683:	DP	Length
Best Local	Similarity	100.0%:	Pred.	No.	0:
Matches	4683:	Conservative	0:	Mismatches	0:
				Indels	0:
				Gaps	0:
QY	1	atgtgatactgctctctcttctgttccagagcttgcgaaatglltccgttgaagacta	60		
DB	453	atgtgatactgctctctctcttctgttccagagcttgcgaaatglltccgttgaagacta	512		
QY	61	cactccagctctactctgttgcacatgcctctgcgaagagtagtgtttgccagcaccag	120		
DB	513	cactccagctctactctgttgcacatgcctctgcgaagagtagtgtttgccagcaccag	572		
QY	121	ggagactctgtgcccctggcccgcagagagcatccctctgtactctgaagatgtaacct	180		
DB	573	ggagactctgtgcccctggcccgcagagagcatccctctgtactctgaagatgtaacct	632		
QY	181	ggcagcggagagagagatctacagatgtccccggatctccgacagctccaccccaagcact	240		
DB	633	ggcagcggagagagagatctacagatgtccccggatctccgacagctccaccccaagcact	692		
QY	241	cctccaatlltccccctccctctctcttaagcttcagacttcaacttatacattat	300		
DB	693	cctccaatlltccccctccctctctcttaagcttcagacttcaacttatacattat	752		
QY	301	latgtcacagcttgaanaatccttcagaggaanaattagaagctcagatgttccatccaagct	360		
DB	753	latgtcacagcttgaanaatccttcagaggaanaattagaagctcagatgttccatccaagct	812		
QY	361	gttttaacggagacctataraagctcagttgtgagagaccgtaaaacatctgagagcaatgt	420		
DB	813	gttttaacggagacctataraagctcagttgtgagagaccgtaaaacatctgagagcaatgt	872		

OY	421	gcggcttccaagtgatcattaccctccctccggtggaggggtacatcactgtggtctcaatg	480
Db	873	ggcgctttccaagtgatcattaccctccctccggtggaggggtacatcactgtggtctcaatg	932
OY	481	gagaaatacacttttccaattgtcccaagatcctaatttccatccacatccacagggagcc	540
Db	933	gagaaatacactgtttccacttgtctccagatccttaattccatccacatccacagggagcc	992
OY	541	tgttatattaaagatgacagaatggaatgtaattgtataactaccgctgtcatcaagcg	600
Db	993	tgttatattaaagatgatacgaatggaatggaattgtataactaccgctgtcatcaagcg	1052
OY	601	catcgaatataccggaagagatgaaatgacgaacagcgcccaacttttgtatacgaacca	660
Db	1053	catcgaatatacccggaagagatgaaatggaagcagacgcccgaacttttgtatacgaacca	1112
OY	661	ggcaactatcgcccatccatccatctcgtgaattgtgaacatctgcgaagacgaatgctggcag	720
Db	1113	ggcaactatcgcccatccatccatctcgtgaattgtgaacatctgcgaagacgaatgctggcag	1172
OY	721	cggtgtgagcttgccttctgtatgaatggtcttcgggaccccttgagccagattaccgtctgctaag	780
Db	1173	cggtgtgagctgtcccttctgtatgaatggtcttcgggaccccttgagccagattaccgtctgctaag	1232
OY	781	gacaaacatgcccctcgtgaactcttctcagggaagtttccagaagacccgtgtacgggctgtcat	840
Db	1233	gacaaacatgcccctcgtgaactcttctcagggaagtttccagaagacccgtgtacgggctgtcat	1292
OY	841	gagaaacatctgcctccctctggaactctcaggcaactctgttttggaagtgttccaaagaatacga	900
Db	1293	gagaaacatctgcctccctctggaactctcaggcaactctgttttggaagtgttccaaagaatacga	1352
OY	901	actgtctaaagtgaaatgagctgcctctcgtacagtgaaacacgcccacttgaaagcccatcagctcc	960
Db	1353	actgtctaaagtgaaatgagctgcctctcgtacagtgaaacacgcccacttgaaagcccatcagctcc	1412
OY	961	aggaaatgtttaaatacgaactgtggtatccagaatttcccttgccttgacggttgaaaggagact	1020
Db	1413	aggaaatgtttaaatacgaactgtggtatccagaatttcccttgccttgacggttgaaaggagact	1472
OY	1021	gaggaacccggaacatctccctctcgtacccgaatggtgtaaatctctcaaccctcgtgaaaaatgtg	1080
Db	1473	gaggaacccggaacatctccctctcgtacccgaatggtgtaaatctctcaaccctcgtgaaaaatgtg	1532
OY	1081	aaagatctatcagaatctcaatctatcagaatctctatgaatggtctcgaatgtctgaacgtgacggg	1140
Db	1533	aaagatctatcagaatctcaatctatcagaatctctatgaatggtctcgaatgtctgaacgtgacggg	1592
OY	1141	gaggaatctcgaatctctctgtgtcgaatgaaatgtgtccgcctcaagactatagtccaggtg	1200
Db	1593	gaggaatctcgaatgtcctctgtgtcgaatgaaatgtgtccgcctcaagactatagtccaggtg	1652
OY	1201	gtcctctgaagatggaacatctcccaaatctattctgccttaagtgaagaagtgtgtgattcca	1260
Db	1653	gtcctctgaagatggaacatctcccaaatctattctgccttaagtgaagaagtgtgtgattcca	1712
OY	1261	ggaagagctcgtttctccctatagtgaaggtgaagaagacacttttgcacagatccaggtg	1320
Db	1713	ggaagagctcgtttctccctatagtgaaggtgaagaagacacttttgcacagatccaggtg	1772
OY	1321	accctcgtgaatggaatgacccgtatcttccaaggtgtgtaagtctcacccgatccagacgaatgatacag	1380
Db	1773	accctcgtgaatggaatgacccgatcttccaaggtgtgtaagtctctgtacccgatccagacgaatgatacag	1832
OY	1381	tcggagggggaacgtgtgtcagcttaccatgaacatctccagctcccaaggtccggggagcgggga	1440
Db	1833	tcggagggggaacgtgtgtcagcttaccatgaacatctccagctcccaaggtccggggagcgggga	1892
OY	1441	gtctacccctcgaactgtccaacaacatctcgggggagctgtcctgttacaggtctgaaataaac	1500
Db	1893	gtctacccgtgtcactgtccaacaacatctcgggggagctgtcctgtgtacccaggtctgaaataaac	1952
OY	1501	gtataagtgctctgaagcatcttgcgaacaatgaanaacatccagacgaatgacgtgac	1560





[illegible]

Db	1533	aggtctacagtgatccacacacgaacacctatattatgatctacatagtcacaaagtgcagcgg	1592
Oy	1141	ggcgataaccagtgctcttgatgcgcgaaggacacgtctccgcctcaagactatgtcaggtg	1200
Db	1593	ggcgataaccagtgctcttgatgcgcgaaggacacgtctccgcctcaagactatgtcaggtg	1652
Oy	1201	gtctctgaagaatgtgaactcccaaatattctctgccttatgtgaaaagtggttgatcca	1260
Db	1653	gtctctgaagaatgtgaactcccaaatattctctgccttatgtgaaaagtggttgatcca	1712
Oy	1261	gcagagcgcgtcttcccttatattgcaacgtgaaaggaaacacctctgtccaccgactcagttg	1320
Db	1713	gcagagcgcgtcttcccttatattgcaacgtgaaaggaaacacctctgtccaccgactcagttg	1772
Oy	1321	acctctgagatgagatgacctgattctcaagggtgagcagtcacggcatcagccagatgatcag	1380
Db	1773	acctctgagatgagatgacctgattctcaagggtgagcagtcacggcatcagccagatgatcag	1832
Oy	1381	tcggaagtgagacgtggtccttctatgcaacgtgaaagctccaccgctccaccaggtcccgagcgggga	1440
Db	1833	tcggaagtgagacgtggtccttctatgcaacgtgaaagctccaccgctccaccaggtcccgagcgggga	1892
Oy	1441	gtctctatctctgagatgacctgaactccgacgagagatctgtctctgacacaggtctcgataaac	1500
Db	1893	gtctctatctctgagatgacctgaactccgacgagagatctgtctctgacacaggtctcgataaac	1952
Oy	1501	gtaagagagctctgcaagatctcgaaccaatggaataacatccagacgaatagcagcagcgagac	1560
Db	1953	gtaagagagctctgcaagatctcgaaccaatggaataacatccagacgaatagcagcagcgagac	2012
Oy	1561	acatacaatctaacgtgctgagatgtgatctgatactccgatactactccatataatgtgtacaagaac	1620
Db	2013	acatacaatctaacgtgctgagatgtgatctgatactccgatactactccatataatgtgtacaagaac	2072
Oy	1621	tctaaccttgccttcccttctaacatccgcgcgaagtgcatcttgagaataatggaactctttaa	1680
Db	2073	tctaaccttgccttcccttctaacatccgcgcgaagtgcatcttgagaataatggaactctttaa	2132
Oy	1681	cttctagaatgagcaaaagagagtgagacgaggggagttacacgtgcaacgvtgtgttcaa	1740
Db	2133	cttctagaatgagcaaaagagagtgagacgaggggagttacacgtgcaacgvtgtgttcaa	2192
Oy	1741	ccacaactctccaccacgacgagtcgtccacgtgacgttgagaaattccgcgctctctatacaa	1800
Db	2193	ccacaactctccaccacgacgagtcgtccacgtgacgttgagaaattccgcgctctctatacaa	2252
Oy	1801	ccctcttgcattctcgaagatctctcgaattgagatgagcagctctcaacgtgtgtgtgtc	1860
Db	2253	ccctcttgcattctcgaagatctctcgaattgagatgagcagctctcaacgtgtgtgtgtc	2312
Oy	1861	tcaagagatcttacctatcacaatctaccttgatgagagagatgagccgacacccctctggagac	1920
Db	2313	tcaagagatcttacctatcacaatctaccttgatgagagagatgagccgacacccctctggagac	2372
Oy	1921	cttagaggtgacatgtgacataattgaactctcagacgtcctctgaaagtattccaatctctcg	1980
Db	2373	cttagaggtgacatgtgacataattgaactctcagacgtcctctgaaagtattccaatctctcg	2432
Oy	1981	ctctatgacacaatgagaaatctaacacctgacatagcccggaatgagccgcgcgctgtgagacac	2040
Db	2433	ctctatgacacaatgagaaatctaacacctgacatagcccggaatgagccgcgcgctgtgagacac	2492
Oy	2041	caaaagccaagtgtgtctgcagaggtctctcccaagcttggtgtctcagccaaggacccaagac	2100
Db	2493	caaaagccaagtgtgtctgcagaggtctctcccaagcttggtgtctcagccaaggacccaagac	2552
Oy	2101	ggagattatggcaaaagcagtcacatccatcaattgtctgtcgcgaagggttaccctgtacatcc	2160
Db	2553	ggagattatggcaaaagcagtcacatccatcaattgtctgtcgcgaagggttaccctgtacatcc	2612
Oy	2161	atcgtgtggaagaattctctaaaggtgtcgtgggtctccccaattccacgcaaatgtccctaaat	2220
Db	2613	atcgtgtggaagaattctctaaaggtgtcgtgggtctccccaattccacgcaaatgtccctaaat	2672

[illegible]

QY	3301	tcaccagaagaacatatacaataactctgtgtccaaactctccaagaaagccttgaaatgaatt	3360
Db	3753	tcaccagaagaacatatacaataactctgtgtccaaactctccaagaaagccttgaaatgaatt	3812
QY	3361	ctccagaggttcagaattatttacttgggccaactcaatgacgaagagcctgggttgagatt	3420
Db	3813	ctccagaggtgtcagaattcaattcaactgggccaactctcatgacggagagcctgggttgagatt	3872
QY	3421	aaaacaatctaccaccacaacaagctctcaatctgagatcgtgacgggtctgaaaagtataccaac	3480
Db	3873	aaaacaatctaccaccacaacaagcctctcaatctgagatcgtgacgggtctgaaaagtataccaac	3932
QY	3481	tacagacatctcagatgtctgtcttcaaccgggcagagaacgggggtcagaagatgaaacagatc	3540
Db	3933	tacagacatctcagatgtctgtcttcaaccgggcagagaacgggggtcagaagatgaaacagatc	3992
QY	3541	ttcaaccgggaccaagaagatgtcttcagaagctctccgggggtgtgaaagcagcgcgcc	3600
Db	3993	ttcaaccgggaccaagaagatgtcttcagaagctctccgggggtgtgaaagcagcgcgcc	4052
QY	3601	tcagccccaatgtctctatgtctcctggctctccctctcaagcttgaagcgatcatccga	3660
Db	4053	tcagccccaatgtctctatgtctcctggctctccctctcaagcttgaagcgatcatccga	4112
QY	3661	aaatacaatctatctctgtgtgtccaaacctctccacaagtatcagagagtttgagggcctct	3720
Db	4113	aaatacaatctatctctgtgtgtccaaacctctccacaagtatcagagagtttgagggcctct	4172
QY	3721	cccgactcgtcttccctcaacaatactcccaactgagatgaaatctgtcagtaagcgtctcgg	3780
Db	4173	cccgactcgtcttccctcaacaatactcccaactgagatgaaatctgtcagtaagcgtctcgg	4232
QY	3781	gtgtgtgtgttactctcagatcgtgaagaagagcaacagagtgaaatcatacaagtcagacca	3840
Db	4233	gtgtgtgtgttactctcagatcgtgaagaagagcaacagagtgaaatcatacaagtcagacca	4292
QY	3841	ctagaacaaagctctctgtgaaatctctgaactctagttgggacagttgactctcatctgagatg	3900
Db	4293	ctagaacaaagctctctgtgaaatctctgaactctagttgggacagttgactctcatctgagatg	4352
QY	3901	aaaagacatctgtctgtct	3960
Db	4353	aaaagacatctgtctgtct	4412
QY	3961	aaagacatctctgagacacacccagctctaglaaacgattgtgagcggggggggacatctctaacg	4020
Db	4413	aaagacatctctgagacacacccagctctaglaaacgattgtgagcggggggggacatctctaacg	4472
QY	4021	aaagacatctctgagacacacccagctctaglaaacgattgtgagcggggggggacatctctaacg	4080
Db	4473	aaagacatctctgagacacacccagctctaglaaacgattgtgagcggggggggacatctctaacg	4532
QY	4081	ggccaaataacatctgagacatctgtatgaaatctctctctctctctctctctctctctctctct	4140
Db	4533	ggccaaataacatctgagacatctgtatgaaatctctctctctctctctctctctctctctctct	4592
QY	4141	gatacaatctctgtct	4200
Db	4593	gatacaatctctgtct	4652
QY	4201	ccctgtgacaaatctgt	4260
Db	4653	ccctgtgacaaatctgt	4712
QY	4261	agtgagcagatggggaggttctccaatcaagccccaagcgaagttctctatcgcttggaat	4320
Db	4713	agtgagcagatggggaggttctccaatcaagccccaagcgaagttctctatcgcttggaat	4772
QY	4321	ctcaaatctgtggagactgtgtgtatgaattcacaatcgacagcccaaatctggagttggccaggg	4380
Db	4773	ctcaaatctgtggagactgtgtgtatgaattcacaatcgacagcccaaatctggagttggccaggg	4832
QY	4381	cgatctatgtgaaatcatacaagaagaagactctaggaagaagcccgagttctcaagagg	4440













```
Db 452 cgtcaagcagtttccagggaaacctacacccgtccggttgaggatcaaatgtaacg 511
Oy 411 aggcagatctcgtcttcaagtcattatccctccctcgggtggagcgctacacactgt 470
Db 512 tggcaacgttgcctctcctcaagtcctacccctcctcagtcgaggaataatgttaagct 571
Oy 471 cgtctcagggagaaagaactgttctcactgtctcaggaactagattctcatcacatc 530
Db 572 tgtatcttgggaagaagaacacagctccatccatccagaagaacaggttttattaccta 631
Oy 531 caccgagacgttataatataaagatgtcaagaagaatgtaatgttataactacccgt 590
Db 632 ccacgagcggtctgtacatctctcagctacaggaagagagccctctccactccgtcg 691
Oy 591 catcacgaggtatctgataccagggagacgagcagagcagacgagcagacttltgt 650
Db 692 catcacagacacaaatataagtcgggagaccgagcagacgaatgggagcagcctctctgt 751
Oy 651 atcaagaccaggaactcagcccatccatatactgatagtgttgcacatcgcaaacat 710
Db 752 gacagacacctgtgagtcgactccacacatccctgatagtctcctccatccaggaagtg 811
Oy 711 ggtctggagcagcgttgaggagctgtgcttgaagaagcgtcgcgcagcctgagcagatcag 770
Db 812 ggcgcgacacacagctggagcttgcccgccctcccggtcgcgcctcagctacatcccgcatccg 871
Oy 771 ctggctgaagagacaacatgccccctgtgaacttctcagggaggttcagaagacggtgacgg 830
Db 872 ctggctcagagatgagtcgagcgcctccctcccggtcagcagcgtctgacacagcagatcatcag 931
Oy 831 gctgtcatttgaagaacatctgcgcctcgcagcctcagcagcgtctgtgttgaagtgctcaaa 890
Db 932 gctgacacatcagacatctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 991
Oy 891 cagatcaggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 950
Db 992 cactctcgtctcgcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1051
Oy 951 catcaatccagcagaggttcaaaagcagcgttggtgagcagcagcagcagcagcagcagcagcagc 1010
Db 1052 cctgacacaaagaagcttgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1111
Oy 1011 gacaggaactcaggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1070
Db 1112 gacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1171
Oy 1071 aaaaatgtgagatcacagaggtacacacagcagcagcagcagcagcagcagcagcagcagcagc 1130
Db 1172 cgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1231
Oy 1131 aagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1190
Db 1232 gacgacatcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1291
Oy 1191 tgtcagagtggtctcttgaagaatggaactcccaaatatttctgacttgaagtgaaaggt 1250
Db 1292 tgcattatctgacttgaagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1351
Oy 1251 ggtgagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1310
Db 1352 ggtcaaccccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1411
Oy 1311 gatacagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1370
Db 1412 ggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1471
Oy 1371 gatgatacagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1392
Db 1472 gtacacatgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1493
```

RESULT 6  
AAV31986/c

```
ID AAV31986 standard; cDNA: 898 bp.
XX
AC AAV31986:
XX
DT 28-SEP-1998 (first entry)
XX
DE Mouse Down syndrome-cell adhesion molecule DS-CAM mid cDNA.
XX
KW DS-CAM; Down syndrome-cell adhesion molecule; neural cell;
KW signal transduction; trisomy 21; mental retardation;
KW holoprosencephaly; corpus callosum agenesis;
KW schizencephaly; diagnosis; assay; mouse; ds; ss.
XX
OS Mus sp.
XX
PN W09817795-A1.
XX
PD 30-APR-1998.
XX
PE 23-OCT-1997; 97WO-US19547.
XX
PR 25-OCT-1996; 96US-0029322.
XX
PA (CEDA-) CEDAKS SINAI MEDICAL CENT.
XX
PI korenberg JK.
XX
DR WP1: 1998-271791/24.
XX
PT New isolated Down's Syndrome-cell adhesion molecule - used to
PT develop products for detection, diagnosis and therapy of
PT developmental and neurological abnormalities
XX
PS Claim 2; Page 81; 109pp; English.
XX
CC This cDNA sequence comprises the middle region of a cDNA clone
CC for murine Down syndrome-cell adhesion molecule (DS-CAM), a
CC member of a novel subclass of the Ig superfamily with homology to
CC neural cell adhesion molecules. The 5' region (see AAV31985)
CC and 3' region (see AAV31987) of the clone are also provided. The
CC murine DS-CAM clone was isolated from a C57 Black/6 mouse brain
CC cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31988) as
CC probe. The invention also provides human DS-CAM proteins (see
CC AAV42086-87), as well as expression vectors and primers useful
CC transgenic animals, antisense oligonucleotides, and primers useful
CC for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are
CC associated with developmental and neurological processes. They can
CC be used in e.g. neural prosthetic devices used in entubulation
CC methods of repelling (repelling) damaged or severed peripheral
CC nerves. The products can also be used in detection, diagnosis and
CC therapy of developmental and neurological abnormalities such as
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides
CC are used for inhibiting translation of mRNA.
XX
SQ Sequence 898 bp; 174 A; 222 C; 245 G; 257 T; 0 other:
XX
Query Match 10.7%; Score 502; DB 19; Length 898;
Best Local Similarity 82.3%; Pred. No. 6.5e-136;
Matches 749; Conservative 0; Mismatches 135; Indels 26; Gaps 14;
Oy 2308 aacgatgtgtgacacaga--cgtcagcaagtcagctactcagcgtttaaattctctgca 2365
Db 892 AACGATGTGGGCGCAGAACGTTTCAGCACTGACACAGGAA-----TTCCGG 838
Oy 2366 tgatacatccatccaataactactacccctgcaagcagcagcagcagcagcagcagcagcagcagc 2425
Db 837 CCATGATATACTTATACCCACACACACCTGGCCATCAGGTCGAAGGAAGA-ATGAGCT 779
Oy 2426 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2485
Db 778 GCACAGCCCAATG--GGAGAGCCCATCATGTTCGGT--GGAGAAAGAGACAGCATCATTA 722
```







Dd	1623	gcgcttcacagctgcgtcaagtcagtgagctcagccacaaccctcagtgaatlygtcttggg	1682
Oy	3629	tccccttcacagctcgaagcgatcatccagaagtagctgtaattctgtccaacccct	3688
		+	
Nb	1683	aagcccgacgltcccccaatgatctctggagggtcacagctggtlactagagccclgca	1742
Oy	3689	atccc 3693	
Dd	1743	gcccc 1747	
 RESULT 11			
AAD05447/c	ID	AAD05447 standard; cDNA: 4336 BP.	
XX	AAD05447:		
AC			
XX			
DT	17-JUL-2001	(first entry)	
XX			
CT	Human secreted protein-encoding gene 17 cDN/ clone HHSFB67, SEQ ID NO:69		
KM	Human; secreted protein; proliferative disorder; cancer; tumour;		
KM	fetal abnormality; developmental abnormality; haematopoietic disorder;		
KM	immune system disorder; AIDS; autoimmune disease; Rheumatoid arthritis;		
KM	Inflammation; allergy; neurological disorder; Alzheimer's disease;		
KM	Parkinson's disease; cognitive disorder; schizophrenia; asthma;		
KM	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;		
KM	cardiovascular disorder; angiogenic disorder; kidney disorder;		
KM	gastrointestinal disorder; pregnancy-related disorder;		
KM	endocrine disorder; infection; wound healing; vulnerability;		
KW	cell culture; chemotaxis; food additive; gene therapy;		
XX	binding partner identification; ss.		
OS	Homo sapiens.		
PH	Key	Location/Qualifiers	
FT	CDS	1297..1416	
FT		/*tag- a	
FT		/product- "Human secreted protein"	
FT		/note- "CDS does not include start codon"	
FT		/partial	
FT	sig-peptide	1297..1387	
FT		/**tag- b	
FT	mat-peptide	1388..1413	
FT		/*tag- c	
FT		/product- "Mature human secreted protein"	
PN	WO200134623-A1.		
	17-MAY-2001.		
XX			
PF	01-NOV-2000; 2000WO-US30037.		
XX			
PK	05-NOV-1999; 99US-0163577.		
PR	30-JUN-2000; 2000US-0215137.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Ruben SM, Komatsoulis GA, Moore PA;		
DR	MPi: 2001-316490/33.		
XX	P-PSDB: AAE01604.		
PT	Nucleic acids encoding 29 human secreted polypeptides, useful for		
XX	preventing, diagnosing and/or treating e.g. cancers, Parkinson's		
XX	disease and diabetic retinopathy -		
PS	Claim 1: page 457-459; 535pp: English.		
CC	AAD05389-AAD05473 represent cDNAs corresponding to 29 human secreted		
CC	protein genes, and AAE01546-AAE01630 represent the proteins they encode.		
CC	AAE01631-AAE01660 represent human secreted protein fragments or variants		
CC	The secreted proteins and their genes are useful for preventing,		

treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 29 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis and treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein-encoding cDNA of the invention.

Sequence 4336 BP; 766 A; 1190 C; 1357 G; 1009 T; 14 other;

Query Match	2.08	Score 91.4	DB 22	Length 4336
Best Local Similarity	45.98	Pred. NO. 1.3e-15		
Matches 388	Conservative	0	Mismatches 451	Indels 6
				Gaps 2

QY	2853	acccttcttctcaactatgagatcgcacatgtagccaaagaacccggaattggcagaagcagc	2911
Db	3152	AGCCCTCTACCTCTCTTAAAGTTCGCGATTAAAGCCGACCATGACATTTGGCCACAGCCAGT	3093
QY	2912	ccagatgaaat---gctccatctcagcagcgagagacagctcctctatgtctccactcaag	2968
Db	3092	TCAGGGAATAGTCCGAATATGATCCACCTCTGCAGGCTGCCCCGATTAAGACCCACACA	3033
QY	2969	aagtcacactgagagcctatcatctcagagcaltcaggtcacatggaagctccccaaga	3028
Db	3032	TCCTCTCCCTTCAGCCGACACCCAGCACACCTCCCTCTATCCGATGCGACGCCGACACAG	2973
QY	3029	acacatcttccaaatggaattatctggtggtctatctgaaatgagttacccgaagtcagcagctg	3088
Db	2972	AACTAAATAATTAATCTAATCTCTCTGAGCTCTTGATATCGAATCCCGAGATGCTCTATG	2913
QY	3089	ggaggtgagcttctgaattctgaattatctcaggtctcagccacagcgagagacagtgagtttaca	3148
Db	2912	AAGGATCTAGGCTACCTTTCACGCTTCCTGAGGCACTACACACCCAGGCGCCACATGGGCTGAG-	2852
QY	3149	cccttgagacaactcaatgaattcactcagtcagcgtctggtgagcgctgtlaaccggg	3208
Db	2853	--CTTAACTATCTGTACACAGTACAGAGCGGTACAGATACGAGTACAGCGCTG	2796
QY	3209	ccggcacggggccttctctctcaggaatactaccaccacactctcgagagatgpgccagct	3268
Db	2795	TGGGTGAGGGGCGCCCTCCAGGCCCCCGCAGAGGCTTTTGTGGGGGCGAGTCCCCACAG	2736
QY	3269	acccccccgaaaatgtccaaagcctagcaacatcacagaagaagcatatcaatctcgtg	3328
Db	2735	CAGCACTCTGTAACCTGCTGCTCCACAGCGCCACAGGCCACACAGCTGGAGCGTGAATTGGG	2676
QY	3329	ccacacttccaaagaaagccttgaatggaattccagaggttccagagttacattacaggg	3388
Db	2675	AGCGACCTCCGTGGACAGCCGAATGAGACATCCAGGGGTTACAGATTATTTCTGGG	2616
QY	3389	ccaacctctagcagcgagagctgggtgagatlaaanaacatcacaccacaacagccttaac	3448
Db	2615	AAGCTTCAGCGCGGGAACCTTCCACAGAGGAGTGAAGACGCTTTTCTCGGCTGGAACACGCG	2556



```

DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:12218.
DE Human: primer: detection; diagnosis: antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
PN 07-FEB-2001.
PD 28-JUL-2000: 2000EP-0116126.
PE 29-JUL-1999: 99JP-0248036.
PR 27-AUG-1999: 99JP-0300253.
PR 11-JAN-2000: 2000JP-0118776.
PR 02-MAY-2000: 2000JP-0183767.
PR 09-JUN-2000: 2000JP-0241899.
XX (HELI-) HELIX RES INST.
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX Claim 8: SEQ ID 12218; 2537bp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 2537 BP; 588 A; 815 C; 700 G; 434 T; 0 other;
SO

```

```

Query Match 1.7%; Score 78; DB 22; Length 2537;
Best Local Similarity 45.6%; Pred. No. 7.Se-12;
Matches 313; Conservative 0; Mismatches 370; Indels 3; gaps 1;

```

```

OY 3008 tcaactgaagaagctcccaagaacattgtcaaatgagatattcgggtacccaatag 3067
   ||||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
DB 3 tccgattgacgcgcgcacagacagagaalcaatgattccctcgggtcccgatccc 62
   ||||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
OY 3068 gtaccagagatcacagacactgggttaactccaaltcaatlatcagltgcagacca 3127
   ||||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
DB 63 gataccggagagctgctctatgaagagcagcgttcacgttcgagcattcaaacacc 122
   ||||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
OY 3128 gcgggagacagtgaggtttacacacctggaacaactgataatgattcaactcagtcgacctg 3187

```

```

DB 123 caggggccacatgggctga---gttaactactcttgaacaagaacagcaggtgaagataac 179
   ||||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
OY 3188 tggltcagagcctgttaaccgycgcagcgggctctcttcttaagaaatcatcaccacca 3247
   ||| ||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
DB 180 ggatgagcgtgtacaaacgtctgggtgaaggccctccagccccccggaaggtcttgg 239
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3248 ctctcgaagatctgcccacaaatlaaccrrccgaaatgttcaaaccaatacaatcacag 3307
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 ctggggaagcagtgcccaacagcagcagcagcagcagcagcagcagcagcagcagc 299
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3308 aagacatataatcctcctggttccactttccaaagaagccttgaattgttcacag 3367
   ||| ||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
DB 300 cacagcttgagcgtgtgacttggagccactctcgtctggaacgcagagatggaatcacag 359
   ||| ||| ||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| |||
OY 3368 ggttcagagatcttaacttgagccaaacttcaagcagcagcagcgttggtgagattaaaca 3427
   ||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 gttcacagaattattctcttggaagcagcagcagcagcagcagcagcagcagcagc 419
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3428 tcaaccacacacagcgttcaacttgagcgttggaagcgttgaaagaatcaccaatacagca 3487
   ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
DB 420 tttcccgcttgagaaacagcgttgaaacttgaacttgacttgctatacgcgctcatatgg 479
   ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
OY 3488 tccagatgtgtgcttctcaccctcagcagcagcagcagcagcagcagcagcagcagc 3547
   ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 tcaagctgagcgtcttcaacgcgcagcgttggtgagcgtctcgagcagcagcagcagcagc 539
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3548 ggaaccaagaagatgttccaggtctctcccgcggtgtgaaagcagcagcagcagcagc 3607
   ||||| ||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
DB 540 agaccacagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 599
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3608 ccatgctcttggtgtcgttccctcctcctcctcctcctcctcctcctcctcctcctc 3667
   ||| ||| ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 600 cctcagatgaatgtgtcgttccctcctcctcctcctcctcctcctcctcctcctcctc 659
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3668 ctgtatctgtctcccaaccctatccc 3693
   ||| ||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
DB 660 ggcgtgtgtaagagcccttgcagccccc 685
   ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 14

AA086478 standard; cDNA: 6000 BP.

AA086478;

17-OCT-1995 (first entry)

Human PTP-erb.

PTP-erb: protein tyrosine phosphatase; osteoblast; differentiation;

osteoclast; osteoporosis; bone; cancer; osteosarcoma; ss.

Homo sapiens.

MO9507935-A.

23-MAR-1995.

09-SEP-1994: 94MO-US10166.

14-SEP-1993: 930S-0122032.

(MERI) MERCK &amp; CO INC.

Rodan GA, Rutledge SJ, Schmidt A;

WPI: 1995-131318/17.

P-PSDB: AAR71726.

Protein tyrosine phosphatase protein PTP-erb specifically expressed

in bone cells - modulators of which are used to treat, e.g.

osteoporosis, and prevent and treat bone loss and cancer.





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:29:06 ; Search time 226.99 Seconds  
(without alignments)  
4672.437 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135  
Perfect score: 4683  
Sequence: 1 atgagactacgtctctc.....gctgcgagagaagcagcc 4683

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

al number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/pctodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/pctodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/pctodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/pctodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/pctodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/pctodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529.2	11.3	1493	2 US-08-752-307B-6	Sequence 6, Appl 1
2	66.4	1.4	6000	1 US-08-348-006B-6	Sequence 6, Appl 1
3	66.4	1.4	6000	2 US-08-800-825A-6	Sequence 6, Appl 1
4	66.4	1.4	6000	4 US-09-158-657-6	Sequence 6, Appl 1
5	66.4	1.4	6000	5 PCT-US94-10166-6	Sequence 6, Appl 1
6	47.8	1.0	2600	2 US-08-427-497E-4	Sequence 4, Appl 1
7	47.8	1.0	3189	2 US-08-427-497E-3	Sequence 4, Appl 1
8	47.8	1.0	3774	2 US-08-341-843B-1	Sequence 1, Appl 1
9	47.8	1.0	3774	2 US-08-427-497E-1	Sequence 1, Appl 1
10	47.8	1.0	3774	2 US-08-427-497E-2	Sequence 2, Appl 1
11	45.8	1.0	5690	2 US-08-447-464-2	Sequence 2, Appl 1
12	45.8	1.0	5690	2 US-08-716-679-2	Sequence 2, Appl 1
13	45.2	1.0	4608	4 US-09-041-886-24	Sequence 24, Appl 1
14	45.2	1.0	4608	5 PCT-US94-05277-14	Sequence 1, Appl 1
15	44.6	1.0	4843	3 US-08-986-485-1	Sequence 1, Appl 1
16	42.8	0.9	4975	2 US-08-249-687C-1	Sequence 1, Appl 1
17	42.8	0.9	4989	2 US-08-666-392A-1	Sequence 3, Appl 1
18	42.8	0.9	4989	2 US-08-625-819-1	Sequence 1, Appl 1
19	42.8	0.9	4989	2 US-08-755-558-4	Sequence 4, Appl 1
20	42.8	0.9	4989	3 US-08-746-559A-1	Sequence 1, Appl 1
21	42.8	0.9	4989	3 US-08-880-313A-9	Sequence 9, Appl 1
22	42.8	0.9	4989	4 US-09-199-926-3	Sequence 3, Appl 1
23	42.8	0.9	4993	3 US-08-746-559A-3	Sequence 3, Appl 1
24	42.2	0.9	3973	1 US-08-036-210-21	Sequence 21, Appl 1
25	42.2	0.9	3973	2 US-08-449-609-21	Sequence 21, Appl 1
26	41	0.9	1581	5 PCT-US95-08493-14	Sequence 14, Appl 1
27	41	0.9	2580	5 PCT-US95-08493-18	Sequence 18, Appl 1

28	41	0.9	2604	5 PCT-US95-08493-20	Sequence 20, Appl 1
29	40.8	0.9	7218	1 US-08-232-463-14	Sequence 5, Appl 1
30	40.6	0.9	3360	1 US-08-408-093-5	Sequence 5, Appl 1
31	40.6	0.9	3360	1 US-08-408-420A-5	Sequence 5, Appl 1
32	40.6	0.9	3360	1 US-08-714-901-5	Sequence 5, Appl 1
33	40.6	0.9	3360	3 US-08-040-741-5	Sequence 10, Appl 1
34	39.2	0.8	2962	2 US-08-449-645A-10	Sequence 10, Appl 1
35	39.2	0.8	2962	2 US-08-702-367A-10	Sequence 10, Appl 1
36	39.2	0.8	2962	5 PCT-US95-04681-10	Sequence 10, Appl 1
37	39	0.8	289	4 US-09-007-005-17	Sequence 17, Appl 1
38	39	0.8	289	4 US-09-244-796-17	Sequence 24, Appl 1
39	38	0.8	1004	4 US-09-383-586-24	Sequence 24, Appl 1
40	36.8	0.8	4149	2 US-08-737-715-1	Sequence 5, Appl 1
41	36.6	0.8	1794	2 US-08-427-497E-5	Sequence 2, Appl 1
42	36.2	0.8	2869	1 US-08-374-834-2	Sequence 2, Appl 1
43	36.2	0.8	2869	1 US-08-644-271-2	Sequence 2, Appl 1
44	35	0.7	564	1 US-08-750-532-11	Sequence 11, Appl 1
45	35	0.7	564	4 US-08-894-818B-14	Sequence 14, Appl 1

## ALIGNMENTS

RESULT 1  
US-08-752-307B-6  
; Sequence 6, Application US/08752307B  
; Patent No. 5952171  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; APPLICANT: Gearing, David P.  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES  
; TITLE OF INVENTION: ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,307B  
; FILING DATE: 19-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McKeljohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 09404/020001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1493 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 99...1493  
; US-08-752-307B-6















QY	3472	TATAGCTTCCTACCACTCGAGAGGTGACAGCCCTTTAAAGGGGCGAGGATTCGGGGCCCCCAGC	2722
QY	3532	gagcagatctctaccccgagacaaagaagatbgtctcaggtctctcccgcgaggtgtgtgaagca	3591
DB	2682	TATAGCTTCCTACCACTCGAGAGGTGACAGCCCTTTAAAGGGGCGAGGATTCGGGGCCCCCAGC	2722
QY	3592	gcgagcgagcctcagcctccatagtgcttctgtctcgtgctctcccccctcaagctgaagcgc	3651
DB	2779	GAGTGCCAGTTCGAACACACACCTCGTCTGCTGCGCTGGACGCCCCCACTACAGCTACAAACGCC	2838
QY	3652	atcaltccgaagtacacatgcatlctcgtctcccaaccctat	3690
DB	2839	GTGCTACACCGCTACGTCGTCCTCTCAAGACCCCTTCGGAT	2877

RESULT 11  
US-08-447-464-2  
Sequence 2, Application US/08447464  
Patent No. 5840842

```

1 GENERAL INFORMATION:
2 APPLICANT: Schlessinger, Joseph
3 APPLICANT: Yan, Hai
4 TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
5 TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
6 NUMBER OF SEQUENCES: 12
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Pennie & Edmonds
9 STREET: 1155 Avenue of the Americas
10 CITY: New York
11 STATE: New York
12 COUNTRY: U.S.A.
13 ZIP: 10036-2711
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/447,464
22 FILING DATE: 24-MAY-1995
23
24 CLASSIFICATION: 435
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 08/130,570
27 FILING DATE: 01-OCT-1993
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Mistrock, S. Leslie
30 REGISTRATION NUMBER: 18,872
31 REFERENCE/DOCKET NUMBER: 7683-043
32
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: 212-790-9090
35 TELEFAX: 212-869-8864/9741
36
37 TELEX: 66141 PENNIE
38 INFORMATION FOR SEQ ID NO: 2:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 5690 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: unknown
43 TOPOLOGY: unknown
44 MOLECULE TYPE: DNA (genomic)
45 FEATURE:
46 NAME/KEY: CDS
47 LOCATION: 833..5338
48
49 US-08-447-464-2

```

```

Query Match      1 0%; Score 45.8; DB 2; Length 5690;
Best Local Similarity 51.7%; Pred. No. 0.0044;
Matches 104; Conservative 0; Mismatches 97; Indels 0; Gaps 0.
QY ^3455 tggacgagcgtcgtgaaggtacaccactatagcgtcctcagttgctggtccttcacccgcagc 3514
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Accession	Sequence	Position
Dd	TGGCACCCTGCTAGAGAGATGAGACCTAACACTGTGAGAGTGTCTGGCTTCACATCGGTGG	2315
Qy	gaaacgagagctcagag, agcagatctcaccggaaccaagaagatgttcaggtctc	3574
Dd	GGCATGGGACCACTGTGACCCCATCAGTCCAGTCAAGACCCAGAGGAGTGGCCGGCCAGC	2375
Qy	ccgagagctgaaagacagcagcgccctcagctcgaatgtcttctgtctcgtgctcc	3634
Dd	GCATGAACTTCTCCGCTGAACTCCAGATCGAGACCAAGCATTTGGCTCTCTGTGAGTGCAC	2435
Qy	ctctcaagctgaacgagatca	3655
Dd	CACCGTAAAGAGATGTTCATTAA	2456

RESULT 12  
US-08-716-679-2  
: Sequence 2, Application 1 /08716679

```

1 PATENT NO. 0846800
2 GENERAL INFORMATION:
3 APPLICANT: Schlessinger, Joseph
4 APPLICANT: Yan, Hai
5 TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
6 TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
7 NUMBER OF SEQUENCES: 12
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Pennie & Edwards
10 STREET: 1155 Avenue of the Americas
11 CITY: New York
12 STATE: New York
13 COUNTRY: U.S.A.
14 ZIP: 10036-2711
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-TMS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/716,679
22 FILING DATE:
23 CLASSIFICATION: 435
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/08/130,570
26 FILING DATE:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Mistrock, S. Leslie
29 REGISTRATION NUMBER: 18,872
30 REFERENCE/BUCKET NUMBER: 7684-043
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 212-790-7090
33 TELEFAX: 212-869 8844/7/41
34 TELEX: 66141 PENNIE
35 INFORMATION FOR SEQ ID NO: 2:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 5690 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: unknown
40 TOPOLOGY: unknown
41 MOLECULE TYPE: DNA (genomic)
42 FEATURE:
43 NAME/KEY: CDS
44 LOCATION: 833..5338
45
46 US-08-716-679-2

```

	Query Match:	1.0%	Score	45.8	DB	2	Length	5690		
	Best Local Similarity	51.7%	Pred.	No.	0.0044					
	Matches	104;	Conservative	0;	Mismatches	97;	Indels	0;	Gaps	0;
OY	3455	tgaacgctcgtcgtagaaagatcaccccaactacacagcalccaggctgcctggccttcaccgccggcag	3514							
Db	2256	TGGCGACCTTGTCTAGAGGATGAGAGCCTCACTCTGTGAAGTGCTGCCTTCATCATTCGTGG	2315							





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:32 ; Search time 6206.58 Seconds  
(without alignments)  
34.627 Million cell updates/sec

Title: US-08-956-991-5

Perfect score: 1 ccaagttccaagagcagg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:  
1: em\_estfun:  
2: em\_esthum:  
3: em\_estln:  
4: em\_estom:  
5: em\_estpl:  
6: em\_estba:  
7: em\_estro:  
8: em\_estov:  
9: em\_hic:  
10: qb\_est1:  
11: qb\_est2:  
12: qb\_hic:  
13: qb\_gss:  
14: em\_gss\_fun:  
15: em\_gss\_hum:  
16: em\_gss\_inv:  
17: em\_gss\_pln:  
18: em\_gss\_pro:  
19: em\_gss\_rtd:  
20: em\_gss\_vrt:  
21: em\_gss\_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	744	11	BF348620 602019924
2	18.4	92.0	538	11	BF348620 602019924
3	17.4	87.0	442	11	BF348620 602019924
4	17.4	87.0	442	11	BF348620 602019924
5	17.4	87.0	442	11	BF348620 602019924
6	16.8	84.0	319	10	AM176762 RCO-CT009
7	16.8	84.0	319	10	AM176762 RCO-CT009
8	16.8	84.0	319	10	AM176762 RCO-CT009
9	16.8	84.0	319	10	AM176762 RCO-CT009
10	16.8	84.0	319	10	AM176762 RCO-CT009
11	16.8	84.0	319	10	AM176762 RCO-CT009
12	16.8	84.0	319	10	AM176762 RCO-CT009

C 13	16.8	84.0	359	11	C84007
C 14	16.8	84.0	380	10	A1347190
C 15	16.8	84.0	380	13	A0996872
C 16	16.8	84.0	393	10	AA200841
C 17	16.8	84.0	400	13	A0894484
C 18	16.8	84.0	444	10	A2363648
C 19	16.8	84.0	449	13	AA188488
C 20	16.8	84.0	467	13	A2229092
C 21	16.8	84.0	469	10	A1636091
C 22	16.8	84.0	475	13	A2363734
C 23	16.8	84.0	482	13	A0150756
C 24	16.8	84.0	502	11	B1426424
C 25	16.8	84.0	526	10	AA160797
C 26	16.8	84.0	583	13	A0307774
C 27	16.8	84.0	603	10	AA4848296
C 28	16.8	84.0	710	11	BG751248
C 29	16.8	84.0	710	13	A2717126
C 30	16.8	84.0	725	10	AU005349
C 31	16.8	84.0	732	11	BG480910
C 32	16.8	84.0	756	13	A2722363
C 33	16.8	84.0	762	11	B1933278
C 34	16.8	84.0	804	11	B145737
C 35	16.8	84.0	826	13	CNS038RC
C 36	16.8	84.0	856	13	A0739885
C 37	16.8	84.0	880	11	B1412208
C 38	16.8	84.0	938	10	BE253944
C 39	16.8	84.0	993	13	CNS056CY
C 40	16.8	84.0	1222	11	BF688244
C 41	16.8	84.0	1224	10	AA694519
C 42	16.4	82.0	312	13	A0339036
C 43	16.4	82.0	374	11	BF391154
C 44	16.4	82.0	376	11	B1295719
C 45	16.4	82.0	390	10	AA770324

#### ALIGNMENTS

RESULT 1  
LOCUS BF348620 114 bp mRNA EST 22-NOV-2000  
DEFINITION 60201:924F1 NCI CGAP, Brn67 Homo sapiens CDNA clone IMAGE:4155508  
5', mRNA sequence.  
ACCESSION BF348620.1 GI:11296215  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgs.nci.nih.gov/  
1 (bases 1 to 744)  
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium (LLNL)  
http://image.llnl.gov  
Plate: LAM9426 row: P column: 05  
High quality sequence stop: 697.  
Location/Qualifiers:  
1. 744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCI CGAP Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"

/lab\_host="DHI08 (T) phage-resistant")  
 /note="Organ: brain; Vector: PCMV-SPORT6; Site: 1; NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI-CGAP Library."  
 BASE COUNT 188 a 196 c 210 g 150 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 744;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 20  
 |||||  
 Db 20 CCAGTCTCAAGAGCAGC 39

RESULT 2  
 BG383716 538 bp mRNA EST 12-MAR-2001  
 DE -NITION 301966 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BG383716  
 VERSION BG383716.1 GI:13308188  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 538)  
 FAHRENKRUG,S.C., FRIEDLING,B.A., ROHRER,G.A., SMITH,T.P.L., CASAS,E.,  
 STONE,R.T., HEATON,M.P., GROSSE,W.M., BENNETT,G.A., LAEGFELD,W.W.,  
 AND KEELE,J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT USDA, ARS, US Meat Animal Research Center  
 Contact: Smith TPL  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -minscore 18  
 and -minmatch 12 options.

PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCACAGCAGCAGC  
 Plate: 88 row: K column: 9  
 Seq primer: ATTAGTGACACTATAG.  
 Location/Qualifiers

FEATURES  
 source 1..538  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9623"  
 /clone\_id="MARC 1P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DHI08"  
 /note="Vector: PCMV-SPORT6; Site: 1; XhoI; Site 2: XhoI;  
 Library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."

BASE COUNT 98 a 150 c 120 g 170 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 538;  
 Best Local Similarity 95.0%; Pred. No. 1.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 20  
 |||||  
 Db 178 CCAGTCTCAAGAGCAGC 197

## RESULT 3

BG348199 442 bp mRNA EST 18-APR-2001  
 LOCUS  
 DEFINITION dg39c10.y1 Xenopus laevis gastrula non normalized Xenopus laevis  
 cDNA clone IMAGE:3751290 5', mRNA sequence.

ACCESSION BG348199  
 VERSION BG348199.1 GI:13168625  
 KEYWORDS EST.  
 SOURCE African clawed frog.

## ORGANISM

Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

## REFERENCE

## AUTHORS

1 (bases 1 to 442)  
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,  
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person  
 ,B., Gibbons,M., Harvey,N., Riller,E., Jackson,Y., McCann,K.,  
 Waterston,R. and Wilson,R.  
 Washu Xenopus EST project, 1999  
 Unpublished (1999)  
 Other-ESTs: dg39c10.x1  
 Contact: Sandy Clifton, Ph.D.  
 Washu Xenopus EST project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

## TITLE

## JOURNAL

## COMMENT

Library constructed by Bruce Blumberg  
 Library normalized by Jihwan Song  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: Xenopus clone distribution information for  
 this library can be found through Research Genetics, visit their  
 web page at: <http://www.resgen.com/> Please reference the id listed  
 below when ordering this clone: Source lab clone id - xlmng011e19

Seq primer: -40RP from Gibco  
 High quality sequence stop: 439.  
 Location/Qualifiers

## FEATURES

## source

1..442  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8155"  
 /clone\_id="IMAGE:3751290"  
 /clone\_id="Xenopus laevis gastrula non normalized"  
 /tissue\_type="gastrula (stages 10.5, 11.5 mixed)"  
 /lab\_host="Top-10 F"  
 /note="Vector: pBluescript SK-; Site: 1; EcoRI; Site 2:  
 XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal  
 parts from stage 10.5 and stage 11.5 gastrulae).  
 EcoRI-XhoI cut cDNA was then ligated into Unizap-XR  
 (Stratagene) with EcoRI at the 5' end and XhoI at the  
 3' end. SS-library phagemids were prepared by mass excision  
 from the original library and normalized by hybridization  
 to biotinylated driver (prepared from the same library by  
 PCR) to Cot-omega of 11. After removal of hybrids and  
 excess driver by streptavidin sepharose chromatography,  
 the ss-phagemids were made double stranded and  
 electroporated into Top-10 F'. Original library  
 construction by Bruce Blumberg (Cho et al. 1991 Cell 67,  
 1111-1120). Note: This is a Xenopus Gene Collection (XGC)  
 library."

BASE COUNT 130 a 105 c 90 g 117 t  
 ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 442;  
 Best Local Similarity 94.7%; Pred. No. 3.8e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 19  
 |||||  
 Db 54 CCATTCTCAAGAGCAGC 72

RESULT 4  
 BI102963/C 739 bp mRNA EST 26-JUN-2001  
 DEFINITION 60288573p1 NCI\_CGAP\_Kid4 Mus musculus cDNA clone IMAGE:5043738  
 5', mRNA sequence.  
 ACCESSION BI102963  
 VERSION BI102963.1 GI:14553856  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-f@mail.nih.gov](mailto:cgaps-f@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LAM1120 row: a column: 19  
 High quality sequence stop: 737.

FEATURES  
 source  
 1..739  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_lib="IMAGE:5043738"  
 /clone\_lib="NCI\_CGAP\_Kid4"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: kidney; Vector: pCMV-Sport6; Site: 1; NCI;  
 Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library. 1"  
 Technology: Note: this is a NCI\_CGAP Library. 1"

BASE COUNT 199 a 201 c 168 g 170 t 1 others  
 ORIGIN  
 Query Match 87.0%; Score 17.4; DB 11; Length 739;  
 Best Local Similarity 94.7%; Pred. No. 3.9e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 cagttctcaaaagagcagc 20  
 ||||||||||||||||  
 202 CAGTCTCAAGACGACATG 184

RESULT 5  
 CDS025H0 832 bp DNA GSS 12-MAY-2000  
 LOCUS Tetraodon nigroviridis genome survey sequence puc-ori end of clone  
 DEFINITION 23616 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.  
 ACCESSION AL182061  
 VERSION AL182061.1 GI:7820152  
 KEYWORDS GSS; genome survey sequence.  
 SOURCE Tetraodon nigroviridis.  
 ORGANISM Tetraodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Crania a; Vertebrata; Euteostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.  
 1 (bases 1 to 832)  
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,  
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Betnol,A. and  
 Weissenbach,J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 832)  
 AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,  
 Bernot,A., Fitzames,C., Wincker,P., Brotier,P., Quetier,F.,  
 Saurin,W. and Weissenbach,J.

TITLE Human gene number estimate provided by genome wide analysis using  
 Tetraodon nigroviridis DNA sequence  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 832)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetraodon nigroviridis  
 genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
 source  
 1..832  
 Location/Qualifiers  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="23616"  
 /clone\_lib="G"  
 /note="Genoscope sequence ID : C0AG236DE08SP1-end ;  
 puc-ori"

BASE COUNT 236 a 221 c 168 g 202 t 5 others  
 ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 832;  
 Best Local Similarity 94.7%; Pred. No. 4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ccagttctcaaaagagcagc 19  
 ||||| |||||||||  
 Db 362 CCACTCTCAAGACGACAGC 344

RESULT 6  
 AM176762 177 bp mRNA EST 16-NOV-1999  
 LOCUS RC0-CT0094-200899-002-B12 CT0094 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AM176762  
 ACCESSION AM176762  
 VERSION AM176762.1 GI:6442799  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.  
 TITLE The FAPESP/LICR Human Cancer Genome Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?lib=RC0&t2=RC0-CT0094-200899-002-B12&t3=1999-08-20&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 177.

FEATURES  
 source

1..177  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="CT0094"  
 /dev\_stage="Adult"  
 /note="Organ: colon; Vector: puc18; Site: 1; SmaI; Site: 2;



ACCESSION AI927023.1 GI:5662987  
 VERSION EST.  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens  
 human.  
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 343)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbsf@mail.nih.gov](mailto:cgapbsf@mail.nih.gov)  
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.  
 Emmerit-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 444 Std Error: 0.00  
 Seq primer: -40UP from Gibco.  
 Location/Qualifiers  
 1..343  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2462290"  
 /clone\_lib="NCI-CCAP\_Kid11"  
 /lab\_host="DH10B"  
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
 Plasmid DNA from the normalized library NCI-CCAP\_Kid3 was  
 prepared, and ss circles were made in vitro. Following HAP  
 purification, this DNA was used as tracer in a subtractive  
 hybridization reaction. The driver was PCR-amplified cDNAs  
 from a pool of 5,000 clones made from the same library  
 (clones 132376-132391, 145607-1456775, and  
 1500552-1502855). Subtraction by Bento Soares and M.  
 Fatima Bonaudo."  
 BASE COUNT 87 a 78 c 97 g 80 t 1 others  
 ORIGIN  
 Query Match 84.0%; Score 16.8; DB 10; Length 343;  
 Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 ccagttctcaagagagag 20  
 ||||| ||||| ||||| |||||  
 ||||| ||||| ||||| |||||  
 Db 136 CCACTCCCAAGACGACG 155  
 RESULT 10  
 LOCUS AW780736 350 bp mRNA EST 12-MAY-2000  
 DEFINITION s184f06.y1 Gm-cl037 Glycine max CDNA GENOME SYSTEMS CLONE ID:  
 Gm-cl037-732.5' Similar to TR:096305 Q96305 SIMILAR TO GLUTAREDOXIN  
 ENCODED BY GENBANK ACCESSION NUMBER Z49699 ;, mRNA sequence.  
 ACCESSION AW780736  
 VERSION AW780736.1 GI:7795411  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 350)  
 Shoemaker,R., Keln,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna  
 A., Holla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 JOURNAL  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@wustl.edu](mailto:est@wustl.edu)  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134 For further information  
 call: (800) 430-0040 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
 427-3424 or contact: [clones@genomesystems.com](mailto:clones@genomesystems.com) or  
[info@genomesystems.com](http://info@genomesystems.com) web site: [www.genomesystems.com](http://www.genomesystems.com)  
 putative full length read  
 vector to vector length is 351.  
 Location/Qualifiers  
 1..350  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl037-732"  
 /clone\_lib="Gm-cl037"  
 /tissue\_type="Fully expanded leaves of greenhouse grown  
 plants."  
 /dev\_stage="2 week old"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI; This  
 cDNA library was constructed from mRNA isolated from fully  
 expanded leaves of greenhouse grown plants that were 2  
 weeks old. The library was prepared using the Life  
 Technologies Psperscript CDNA library construction kit.  
 Complementary DNA was synthesized from mRNA using a  
 poly(dT) sequence with a NotI restriction site. SalI  
 linkers adapters were ligated to the blunt ended cDNA  
 fragments followed by NotI digestion. The cDNA fragments  
 were directionally cloned into the NotI-SalI restriction  
 site of the pSPORT1 vector. The ligated cDNA fragments  
 were transformed into E.coli Electro-Max DH10B host  
 cells. This library was constructed in the laboratory of  
 Dr. Jila Vodkin by Anu Khanna at the University of  
 Illinois at Urbana-Champaign. email: [l-vodkin@uiuc.edu](mailto:l-vodkin@uiuc.edu)"  
 BASE COUNT 93 a 46 c 65 g 146 t  
 ORIGIN  
 Query Match 84.0%; Score 16.8; DB 10; Length 350;  
 Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 ccagttctcaagagagag 20  
 ||||| ||||| ||||| |||||  
 ||||| ||||| ||||| |||||  
 Db 102 CCAATTCTCAAGACGACG 121  
 QY 1 ccagttctcaagagagag 20  
 ||||| ||||| ||||| |||||  
 ||||| ||||| ||||| |||||  
 RESULT 11  
 LOCUS N24917 356 bp mRNA EST 28-DEC-1995  
 DEFINITION y01h12.s1 Soares melanocyte 2NBHM Homo sapiens CDNA clone  
 IMAGE:270023.3', mRNA sequence.  
 ACCESSION N24917  
 VERSION N24917.1 GI:1139067  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 356)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,D.,  
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston



TITLE  
JOURNAL  
COMMENT

R. Williamson, A., Wohlmann, P. and Wilson, R.  
The Washu-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
High quality sequence stops: 254  
Source: IMAGE Consortium, LNLN.  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: ml3 -40 forward  
High quality sequence stop: 254.

## FEATURES

source

1. 356  
/organism="Homo sapiens"  
/db\_xref="GDB:3879665"  
/db\_xref="taxon:9606"  
/clone="IMAGE:270023"  
/clone\_lib="Soares melanocyte 2NBHM"  
/sex="Male"  
/tissue="type="melanocyte"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: pRT73D (pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - c1190(dT) primer (5'  
TGTATCAATCTGAGTGGAGCGCGCGCTTTTCTTTTCTT 3'),  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pRT73 vector  
(pharmacia). Library constructed by Bento Soares and  
M. Fatima Bonaldo. RNA from normal foreskin melanocytes  
(FS34) was kindly provided by Dr. Anthony P. Albino."

## BASE COUNT

90 a 81 c 103 g 82 t

## ORIGIN

Query Match 84.0%; Score 16.8; DB 11; Length 356;  
Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccaagttccaagaagcagcag 20  
||||| | ||||| ||||| |||||  
Db 138 CCAAGTCCCAAGAGCAGCAG 157

## A T 12

LOCUS AM015308 358 bp mRNA EST 10-SEP-1999  
DEFINITION UT-H-B10-aap-e-04-0-01.s1 NCLCGAP\_Sub1 Homo sapiens cDNA clone  
IMAGE:2709991.3, mRNA sequence.

ACCESSION AM015308  
VERSION AM015308.1 GI:5864065  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 358)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution Information can be found through the  
I.M.A.C.E. Consortium/LNLN at:

## FEATURES

source

www.bio.lnl.gov/db/rlp/image/image.html  
Seq primer: M13 forward  
POLYA=yes.

Location/Qualifiers  
1. 358

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2709991"  
/clone\_lib="NCLCGAP\_Sub1"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pRT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The  
NCLCGAP\_Sub1 library is a subtracted library derived from  
BL. It constitutes a mixture of 21 normalized or  
subtracted NCLCGAP libraries: NCLCGAP\_C04,  
NCLCGAP\_Pr22, NCLCGAP\_Pr28, NCLCGAP\_C010, NCLCGAP\_C016,  
NCLCGAP\_Kid5, NCLCGAP\_Kid12, NCLCGAP\_Kid3,  
NCLCGAP\_Kid11, NCLCGAP\_Lym2, NCLCGAP\_Br23, NCLCGAP\_Lu5,  
NCLCGAP\_C041, NCLCGAP\_Lu12, NCLCGAP\_GC4, NCLCGAP\_GC6,  
NCLCGAP\_Lu24, NCLCGAP\_Lu19, NCLCGAP\_GC4, NCLCGAP\_C08,  
NCLCGAP\_Brn25. These 21 libraries were pooled and a  
single-stranded DNA preparation of the resulting mixture  
was used as a tracer in a subtractive hybridization with a  
driver whose composition is detailed below: NCLCGAP\_Kid3  
pool 1 LLM 3334-3337, 3682-3683, 3798-3803 (IMAGE  
Clonoids 1322376-1323911, 1456008-1456775, 1500552-1502855  
) NCLCGAP\_Kid5 pool 1 LLM 3338-3342, 3722-3725.  
3776-3778 (IMAGE Clonoids 1323912-1325831,  
1471368-1472903, 1492104-1493255) NCLCGAP\_Lu5 pool 1 LLM  
3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991,  
1520904-1522439) NCLCGAP\_GC4 pool 1 LLM 3164-3167,  
3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,  
1469064-1470983, 1475592-1476743) NCLCGAP\_Pr22 pool 1  
LLM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids  
985608-986759, 1101192-1101928, 1217928-1220615)  
NCLCGAP\_C010 pool 1 LLM 2644-2653, 2871-2872 (IMAGE  
Clonoids 1057416-1061255, 1144584-1145351) The resulting  
subtracted library contained 530,000 recombinants.  
Subtraction was performed as previously described [Bonaldo  
& Lennon & Soares (1996)]: Normalization and Subtraction:  
Two Approaches To Facilitate Gene Discovery. Genome  
Research 6, 791-806.  
TAG LIB-NCLCGAP\_Pr22  
TAG TISSUE-Prostate  
TAG\_SEQ-ANGTC"

BASE COUNT 85 a 79 c 96 g 98 t

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 358;  
Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccaagttccaagaagcagcag 20  
||||| | ||||| ||||| |||||  
Db 156 CCAAGTCCCAAGAGCAGCAGC 175

## RESULT 13

LOCUS C84007 359 bp mRNA EST 28-APR-1999  
DEFINITION C84007 Dictyostelium discoideum SS (H. Urushihara) Dictyostelium  
discoideum cDNA clone SSC227, mRNA sequence.

ACCESSION C84007  
VERSION C84007.1 GI:2723613  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum.  
ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
1 (bases 1 to 359)  
Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,  
Yoshino, R., Mita, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,  
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.



Query Match 84.0%; Score 16.8; DB 13; Length 380;  
 Best Local Similarity 90.0%; Pred. No. 7.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ccagttctcaagagcagc 20  
 |||||  
 Db 123 CCACTTCTCAAGTACCAGG 142

Search completed: March 4, 2002, 02:14:35  
 Job time: 6484 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:21:18 : Search time 578.64 Seconds  
(without alignments)  
29,632 Million cell updates/sec

Title: US-08-956-991-5

Perfect score: 20  
Sequence: 1 ccagttcacaagagcagc 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_1101:\*

1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	19	AAV31983 Human Down syndrome
2	20	100.0	6413	19	AAV31988 Human Down syndrome
3	20	100.0	6604	19	AAV31981 Human Down syndrome
4	18.4	92.0	2173	19	AAV31987 Mouse Down syndrome
5	18.4	92.0	54548	21	AAZ4596 DNA sequence of lh
6	17.4	87.0	2197	21	AAZ89617 A. sydowi afl-SST
7	16.8	84.0	187	16	AAZ22739 Human gene signatu
8	16.4	82.0	12923	10	AAAG0338 Sequence of human
9	15.8	79.0	209	22	AAI25968 Probe #15901 for 9
10	15.8	79.0	209	22	AAI53359 probe #22045 used
11	15.8	79.0	563	22	AAI16773 Probe #6706 for ge

C	12	15.8	79.0	563	22	AAI40280
	13	15.8	79.0	564	22	AAI48672
	14	15.8	79.0	564	22	AAI08987
	15	15.8	79.0	673	22	AAH00627
	16	15.8	79.0	1972	22	AAI35523
	17	15.8	79.0	1972	22	AAI03989
	18	15.8	79.0	2947	22	AAH18697
	19	15.8	79.0	5406	19	AAV52289
	20	15.4	77.0	530	22	AAH13400
	21	15.4	77.0	796	22	AAH05981
	22	15.4	77.0	1665	22	AAH14284
	23	15.4	77.0	4646	20	AAH20577
	24	15.4	77.0	4949	21	AAH76989
	25	15.2	76.0	47	21	AAZ68502
	26	15.2	76.0	167	22	AAH09066
	27	15.2	76.0	180	22	AAI51608
	28	15.2	76.0	351	21	AAH74771
	29	15.2	76.0	478	22	AAI38475
	30	15.2	76.0	750	22	AAI59635
	31	15.2	76.0	798	22	AAI57849
	32	15.2	76.0	892	21	AAA49173
	33	15.2	76.0	1209	20	AAH81466
	34	15.2	76.0	1228	20	AAH81465
	35	15.2	76.0	1229	20	AAH81464
	36	15.2	76.0	1402	21	AAH56896
	37	15.2	76.0	1407	22	AAH42469
	38	15.2	76.0	1408	21	AAH43700
	39	15.2	76.0	1881	22	AAH22645
	40	15.2	76.0	1881	22	AAH27735
	41	15.2	76.0	1930	22	AAH52881
	42	15.2	76.0	1996	22	AAH15280
	43	15.2	76.0	2303	22	AAH30955
	44	15.2	76.0	2336	22	AAH16558
	45	15.2	76.0	2652	21	AAH61272

#### ALIGNMENTS

RESULT 1

AAV31983

ID AAV31983 standard. cDNA: 20 BP.

AC AAV31983:

XX 28-SEP-1998 (first entry)

XX DT

XX DE Human Down syndrome-cell adhesion molecule US-CAM primer B9-131F.

XX AC

XX DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;

KW signal transduction; trisomy 21; mental retardation;

KW holoprosencephaly; corpus callosum agenesis;

KW schizencephaly; diagnosis; assay; human; PCR; primer; ss.

XX OS Synthetic.

OS Homo sapiens.

XX OS

XX PN WC9817795-A1.

XX PD

XX 30-APR-1998.

XX PF

XX 23-OCT-1997. 97MO-US19547.

XX PR

XX 25-OCT-1996. 96US-0029322.

XX PA

XX (CEDA-) CEUAKS SINAI MEDICAL CENT.

XX PI

XX Korenberg JR.

XX DR

XX WPI: 1998-271791/24.

XX PT

PT New isolated Down's syndrome-cell adhesion molecule - used to  
develop products for detection, diagnosis and therapy of

```

PT developmental and neurological abnormalities
XX
XX Example 5: Page 79; 109pp: English.
PS
XX Forward primer B9-131F and reverse primer B9-131R (see AAV31984)
CC were used in RT-PCR assays of cDNA libraries of various human
CC tissues to determine human Down syndrome-cell adhesion molecule
CC DS-CAM (see AAV31981 and AAV31988) expression. The results
CC demonstrated expression of human DS-CAM mRNA in foetal and adult
CC brain, and foetal kidney. In addition, a breast carcinoma cell
CC line showed expression of human DS-CAM mRNA. DS-CAM polypeptides
CC (see AAV42086-87) are associated with developmental and neurological
CC processes. The polypeptides and nucleic acids are used to develop
CC products for the detection, diagnosis and therapy of developmental
CC and neurological abnormalities.
XX
XX Sequence 20 BP: 6 A; 5 C; 6 G; 3 T; 0 other;
SQ
} Query Match 100.0%; Score 20; DB 19; Length 20;
{ Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ccaggtctcaagagcagc 20
Db 1 ccaggtctcaagagcagc 20
}
RESULT 2
AAV31988
ID AAV31988 standard; cDNA: 6413 BP.
XX
XX AAV31988:
XX
XX 28-SEP-1998 (first entry)
XX
XX Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.
DE
XX
XX DS-CAM2: Down syndrome-cell adhesion molecule; neural cell;
XX signal transduction; trisomy 21; mental retardation;
XX holoprosencephaly; corpus callosum agenesis;
XX schizencephaly; diagnosis; assay; human; ds; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX FH 453..5168
XX FT /*tag= a
XX FT
XX
XX #09817795-A1.
XX
XX 30-APR-1998.
XX
XX 23-OCT-1997; 97WO-US19547.
XX
XX 25-OCT-1996; 96US-0029322.
XX
XX (CEDA-) CEVARS SINAI MEDICAL CENT.
XX
XX Korenberg JK:
XX
XX WPI: 1998-271791/24.
XX P-PSDB: AAM42087.
XX
XX New isolated Down's Syndrome-cell adhesion molecule - used to
XX develop products for detection, diagnosis and therapy of
XX developmental and neurological abnormalities
XX
XX Claim 3: Page 83-90; 109pp: English.
XX
XX This cDNA clone codes for Down syndrome-cell adhesion molecule
XX DS-CAM2 (see AAM42087), an extracellular, soluble protein belonging
XX to a novel subclass of the Ig superfamily with highest homology to

```

```

CC neural cell adhesion molecules. A modified direct cDNA selection
CC technique was applied to bacterial and pl artificial chromosomes
CC between ETS2 and MX1 by using cDNA from trisomy 21 human foetal
CC brain. A unique cDNA fragment, designated E51 (see AAV31982), was
CC obtained and used to screen a trisomy 21 human foetal brain (14-wk
CC gestation) cDNA library. Further clones were obtained by exon
CC trapping, and the complete DS-CAM2 cDNA sequence was identified.
CC DS-CAM2 is a splice variant of membrane-bound DS-CAM1 (see AAM42086)
CC and is deleted for the entire transmembrane domain. The DS-CAM
CC gene spans 900-1200 kb of genomic DNA and maps at chromosome
CC 21q22.2-22.3. The invention also provides murine DS-CAM partial
CC sequences (see AAV31985-87), expression vectors and host cells,
CC transgenic animals, antisense oligonucleotides, and primers useful
CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are
CC associated with developmental and neurological processes. They can
CC be used in e.g. neural prosthetic devices used in entubulation
CC methods of repairing (regenerating) damaged or severed peripheral
CC nerves. The products can also be used in detection, diagnosis and
CC therapy of developmental and neurological abnormalities such as
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides
CC are used for inhibiting translation of mRNA.
XX
XX SQ
} Query Match 100.0%; Score 20; DB 19; Length 6413;
{ Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ccaggtctcaagagcagc 20
Db 4877 ccaggtctcaagagcagc 4896
}
RESULT 3
AAV31981
ID AAV31981 standard; cDNA: 6604 BP.
XX
XX AAV31981:
XX
XX 28-SEP-1998 (first entry)
XX
XX Human Down syndrome-cell adhesion molecule DS-CAM1 cDNA.
DE
XX
XX DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;
XX signal transduction; trisomy 21; mental retardation;
XX holoprosencephaly; corpus callosum agenesis;
XX schizencephaly; diagnosis; assay; human; ds; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX FH 453..6185
XX FT /*tag= a
XX FT sig_peptide 452..521
XX FT mat_peptide 522..6162
XX FT /*tag= c
XX FT
XX #09817795-A1.
XX
XX 30-APR-1998.
XX
XX 23-OCT-1997; 97WO-US19547.
XX
XX 25-OCT-1996; 96US-0029322.
XX
XX (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
XX Korenberg JK:
XX
XX WPI: 1998-271791/24.
XX

```

UK P-PSDB: AAM42086.  
 XX New isolated Down's Syndrome-cell adhesion molecule - used to  
 PT develop products for detection, diagnosis and therapy of  
 PT developmental and neurological abnormalities  
 XX  
 PS Claim 3: Page 65-72; 109pp; English.  
 XX  
 CC This cDNA clone codes for Down syndrome-cell adhesion molecule  
 CC DS-CAM1 (see AAM42086), a cell surface glycoprotein belonging to a  
 CC novel subclass of the Ig superfamily with highest homology to  
 CC neural cell adhesion molecules. A modified direct cDNA selection  
 CC technique was applied to bacterial and plasmid artificial chromosomes  
 CC between ET52 and MX1 by using cDNA from trisomy 21 human foetal  
 CC brain. A unique cDNA fragment, designated E31 (see AAV31982), was  
 CC obtained and used to screen a trisomy 21 human foetal brain (14-wk  
 CC gestation) cDNA library. Further clones were obtained by exon  
 CC trapping. The complete DS-CAM1 cDNA sequence was deduced from  
 CC overlapping clones. A splice variant cDNA (see AAV31988), encoding  
 CC non-membrane bound DS-CAM2 (see AAM42087), was also identified. The  
 CC DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome  
 CC 21q22.2-22.3. The invention also provides murine DS-CAM partial  
 CC sequences (see AAV31985-87), expression vectors and host cells,  
 CC transgenic animals, antisense oligonucleotides, and primers useful  
 CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are  
 CC associated with developmental and neurologic processes. They can  
 CC be used in e.g. neural prosthetic devices used in enervation  
 CC methods of repairing (regenerating) damaged or severed peripheral  
 CC nerves. The products can also be used in detection, diagnosis and  
 CC therapy of developmental and neurological abnormalities such as  
 CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
 CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
 CC are used for inhibiting translation of mRNA.  
 XX  
 SQ Sequence 6604 BP; 1673 A; 1827 C; 1768 G; 1336 T; 0 other;  
 QY 1 ccaagttctcaaaagagcagc 20  
 Db 4877 ccaagttctcaaaagagcagc 4896  
 RE:SLIT 4  
 A 1987  
 I AAV31987 standard; cDNA: 2173 BP.  
 AC AAV31987:  
 XX 28-SEP-1998 (first entry)  
 DT  
 XX Mouse Down syndrome-cell adhesion molecule DS-CAM 3' cDNA.  
 DE  
 XX  
 KW DS-CAM: Down syndrome-cell adhesion molecule; neural cell;  
 KW signal transduction; trisomy 21; mental retardation;  
 KW holoprosencephaly; corpus callosum agenesis;  
 KW schizencephaly; diagnosis; assay; mouse; ds; ss.  
 XX  
 OS Mus sp.  
 US  
 XX MO9817795-A1.  
 PN  
 XX 30-APR-1998.  
 PD  
 XX 23-OCT-1997; 97WO-US19547.  
 PF  
 XX 25-OCT-1996; 96US-0029322.  
 PR  
 XX (CEDA-) CEDARS SINAI MEDICAL CENT.  
 PA  
 XX

PI Korenberg JK:  
 XX  
 DR WPI: 1998-271791/24.  
 XX  
 XX New isolated Down's Syndrome-cell adhesion molecule - used to  
 PT develop products for detection, diagnosis and therapy of  
 PT developmental and neurological abnormalities  
 XX  
 PS Claim 2: Page 81-83; 109pp; English.  
 XX  
 CC This cDNA sequence comprises the 3' region of a cDNA clone  
 CC for murine Down syndrome-cell adhesion molecule (DS-CAM), a  
 CC member of a novel subclass of the Ig superfamily with homology to  
 CC neural cell adhesion molecules. The middle region (see AAV31986)  
 CC and 5' region (see AAV31987) of the clone are also provided. The  
 CC murine DS-CAM clone was isolated from a C57 Black/6 mouse brain  
 CC cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31988) as  
 CC probe. The invention also provides human DS-CAM proteins (see  
 CC AAM42086-87), as well as expression vectors and host cells,  
 CC transgenic animals, antisense oligonucleotides, and primers useful  
 CC for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are  
 CC associated with developmental and neurological processes. They can  
 CC be used in e.g. neural prosthetic devices used in enervation  
 CC methods of repairing (regenerating) damaged or severed peripheral  
 CC nerves. The products can also be used in detection, diagnosis and  
 CC therapy of developmental and neurological abnormalities such as  
 CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
 CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
 CC are used for inhibiting translation of mRNA.  
 XX  
 SQ Sequence 2174 BP; 585 A; 598 C; 554 G; 436 T; 0 other;  
 QY 1 ccaagttctcaaaagagcagc 20  
 Db 479 ccaagttctcaaaagagcagc 498  
 RE:SLIT 5  
 AA245596  
 ID AA245596 standard; DNA: 54548 BP.  
 XX  
 AC AA245596:  
 XX 06-APR-2000 (first entry)  
 DT  
 XX DNA sequence of the wild type human PCYGA region of Xq13.  
 DE  
 XX Human; PCYGA region; X chromosome; q13 region; polymorphism;  
 KW mental retardation; autism; depression; bipolar affective disorder;  
 KW hypothyroidism; OPA gene; neuropsychiatric disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FH exon  
 FT 1391..1490  
 FT /tag- a  
 FT /number- 1  
 FT /note- "human OPA exon"  
 FT 1398..1648  
 FT /tag- b  
 FT /note- "human OPA promoter"  
 FT promoter  
 FT complement (1565..1815)  
 FT /tag- c  
 FT /note- "unknown promoter"  
 FT 2008..2112  
 FT /tag- d  
 FT /number- 2  
 FT /note- "human OPA promoter"  
 FT

FT	exon	2321..2512	/*tag= e	FT	/*tag= v	FT	/*tag= w
FT		/number= 3	/note= "human OPA promoter"	FT		FT	/number= 21
FT	exon	2649..2805		FT	exon	FT	/note= "human OPA promoter"
FT		/*tag= f		FT		FT	10330..10757
FT		/number= 4		FT		FT	/*tag= x
FT	exon	3608..3787	/note= "human OPA promoter"	FT	exon	FT	/number= 22
FT		/*tag= g		FT		FT	/note= "human OPA promoter"
FT		/number= 5		FT		FT	10933..11077
FT	exon	3962..4072	/note= "human OPA promoter"	FT		FT	/*tag= y
FT		/*tag= h		FT		FT	/number= 23
FT		/number= 6		FT	exon	FT	/note= "human OPA promoter"
FT		/note= "human OPA promoter"		FT		FT	11235..11355
FT	exon	4197..4451		FT		FT	/*tag= z
FT		/*tag= i		FT		FT	/number= 24
FT		/number= 7		FT		FT	/note= "human OPA promoter"
FT		/note= "human OPA promoter"		FT	exon	FT	11751..12065
FT	exon	4839..4986		FT		FT	/*tag= aa
FT		/*tag= j		FT		FT	/number= 25
FT		/number= 8		FT		FT	/note= "human OPA promoter"
FT	exon	5148..5247	/note= "human OPA promoter"	FT	exon	FT	12317..12492
FT		/*tag= k		FT		FT	/*tag= ab
FT		/number= 9		FT		FT	/number= 26
FT		/note= "human OPA promoter"		FT		FT	/note= "human OPA promoter"
FT	exon	5377..5513		FT	exon	FT	12672..12851
FT		/*tag= l		FT		FT	/*tag= ac
FT		/number= 10		FT		FT	/number= 27
FT		/note= "human OPA promoter"		FT		FT	/note= "human OPA promoter"
FT	exon	5734..5865		FT		FT	14187..14258
FT		/*tag= m		FT		FT	/*tag= ad
FT		/number= 11		FT		FT	/number= 28
FT		/note= "human OPA promoter"		FT	exon	FT	/note= "human OPA promoter"
FT		/*tag= n		FT		FT	14710..14843
FT		/number= 12		FT		FT	/*tag= ae
FT	exon	6798..7027		FT		FT	/number= 29
FT		/*tag= o		FT		FT	/note= "human OPA promoter"
FT		/number= 13		FT	exon	FT	15014..15175
FT		/note= "human OPA promoter"		FT		FT	/*tag= af
FT	exon	7403..7483		FT		FT	/number= 30
FT		/*tag= p		FT		FT	/note= "human OPA promoter"
FT		/number= 14		FT	exon	FT	15487..15598
FT		/note= "human OPA promoter"		FT		FT	/*tag= ag
FT	exon	7615..7784		FT		FT	/number= 31
FT		/*tag= q		FT		FT	/note= "human OPA promoter"
FT		/number= 15		FT	exon	FT	15765..15854
FT		/note= "human OPA promoter"		FT		FT	/*tag= ah
FT	exon	7989..8133		FT		FT	/number= 32
FT		/*tag= r		FT		FT	/note= "human OPA promoter"
FT		/number= 16		FT	exon	FT	16999..17109
FT		/note= "human OPA promoter"		FT		FT	/*tag= ai
FT	exon	8301..8351		FT		FT	/number= 33
FT		/*tag= s		FT		FT	/note= "human OPA promoter"
FT		/number= 17		FT	exon	FT	17351..17492
FT		/note= "human OPA promoter"		FT		FT	/*tag= aj
FT	exon	8674..8792		FT		FT	/number= 34
FT		/*tag= t		FT		FT	/note= "human OPA promoter"
FT		/number= 18		FT	exon	FT	17736..17897
FT		/note= "human OPA promoter"		FT		FT	/*tag= ak
FT	exon	8979..9122		FT		FT	/number= 35
FT		/*tag= u		FT		FT	/note= "human OPA promoter"
FT		/number= 19		FT	exon	FT	18927..19301
FT		/note= "human OPA promoter"		FT		FT	/*tag= al
FT	exon	9606..9769		FT		FT	/number= 36
FT		/*tag= v		FT	exon	FT	/note= "human OPA promoter"
FT		/number= 20		FT		FT	19525..19675
FT	exon	9973..10104		FT		FT	/*tag= am
FT				FT		FT	/number= 37
FT				FT	exon	FT	/note= "human OPA promoter"
FT				FT		FT	19833..20029
FT				FT		FT	/*tag= an
FT				FT		FT	/number= 38
FT				FT		FT	/note= "human OPA promoter"
FT	exon			FT	exon	FT	20207..20281
FT				FT		FT	/*tag= ao

```

FT /number= 39
FT /note= "human OPA promoter"
FT 20372..20590
FT /*tag= ap
FT /number= 40
FT /note= "human OPA promoter"
FT 23283..23505
FT /*tag= aq
FT /number= 41
FT /note= "human OPA promoter"
FT 23878..24018
FT /*tag= at
FT /number= 42
FT /note= "human OPA promoter"
FT 24531..24612
FT /*tag= as
FT /number= 43
FT /note= "human OPA promoter"
FT 24823..25096
FT /*tag= at
FT /number= 44
FT /note= "human OPA promoter"
FT 26123..26275
FT /*tag= au
FT /note= "probable NL-3 promoter"

Query Match      92.0%; Score 18.4; DB 21; Length 54548;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 ccaagttcctcaagaagcagc 20
    ||||| ||||| ||||| |||||
Db 28196 ccaagttcctcaagaagcagc 28215

```

```

RESULT 6
AAZ89617/c
ID AAZ89617 standard; cDNA: 2197 BP.
XX
AC AAZ89617;
XX
DT 23-JUN-2000 (first entry)
XX
DE A. sydowi afl-SST cDNA.
XX
KM Fructosyl transferase; afl-SST; polyfructose; surfactant manufacture;
KM viscosity; sedimentation; food; ss.
XX
A. Aspergillus sydowi.

```

```

A.. Location/Qualifiers
FH Key 21..2069
FT CDS /tag= a
FT /product= "afl-SST"
XX
XX DE19840028-A1.
XX
XX PD 09-MAR-2000.
XX
XX PF 02-SEP-1998; 98DE-1040028.
XX
XX PR 02-SEP-1998; 98DE-1040028.
XX
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX PI Heyer A, Reim J, Wendenburg R;
XX
XX DK WPI: 2000-257741/23.
XX
XX DR P-PSDB: AAT51947.
XX
XX PT Novel DNA encoding Aspergillus fructosyl transferase useful for
XX production of polyfructose which is used to manufacture surfactants and
XX suspension agents
PT

```

```

XX
PS Claim 1b; Page 12-15; 28pp; German.
XX
CC This invention describes a novel nucleic acid molecule (I) encoding a
CC fructosyl transferase isolated from Aspergillus sydowi. The fructosyl
CC transferase is useful for the production of polyfructose. Polyfructose
CC can be used to manufacture surfactants, to increase the viscosity of an
CC aqueous system, as a suspending agent, to accelerate sedimentation and
CC in the formation of complexes or to bind water. The host cell can be
CC used as a food. This sequence encodes the A. sydowi fructosyl transferase
CC designated afl-SST which is described in the method of the invention.
XX
SO Sequence 2197 BP; 515 A; 587 C; 575 G; 520 T; 0 other;

```

```

Query Match      87.0%; Score 17.4; DB 21; Length 2197;
Best Local Similarity 94.7%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 ccaagttcctcaagaagcagc 19
    ||||| ||||| ||||| |||||
Db 728 CCAGTTCCTCAAGAGCAGC 710

```

```

RESULT 7
AAT22739/c
ID AAT22739 standard; cDNA to mRNA: 187 BP.
XX
XX AAT22739;
XX
DT 14-AUG-1996 (first entry)
XX
DE Human gen. signature H0KGS04398.
XX
KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KM cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
XX PN WO9514772-A1.
XX
XX PD 01-JUN-1995.
XX
XX PF 11-NOV-1994; 94W01-100191b.
XX
XX PR 12-NOV-1994; 94JP-045504.
XX
XX PA (MATSU//) MATSUHARA K.
XX
XX PI (OKU//) OKUHO K.
XX
XX PS Matsubara K, Okuho K;
XX
XX WPI: 1995-206931/27.
XX
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX for diagnosis of abnormal cell function, by preparing cDNA that
XX reflects relative abundance of corresp. mRNA in specific human
XX tissues
XX
XX PS Claim 1; Page 1203; 2245pp; Japanese.
XX

```

```

A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-126837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.

```



CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequences) as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types.  
 CC  
 SQ Sequence 187 BP: 48 A; 42 C; 42 G; 49 T; 6 other;

Query Match 84.0%; Score 16.8; DB 16; Length 187;  
 Best Local Similarity 90.0%; Pred. No. 64;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 20  
 ||||| ||||| ||||| |||||  
 DB 46 CCAGTCCCAAGAGCAGG 27

RESULT 8  
 AA90338/C  
 17 AA90338 standard; cDNA: 12923 BP.

AA90338;

29-MAR-1992 (first entry)

Sequence of human muscular dystrophy (MD) cDNA.

Dystrophin: muscular dystrophy; probe; antibody; diagnosis;  
 prenatal; heterozygote; gene therapy; genetic screening;  
 foetal screening; ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 209..12923  
 /\*tag= a

W08906286-A.

13-JUL-1989.

16-DEC-1988; 88WO-US04504.

22-DEC-1987; 87US-0136618.

(CHIL-) CHILDRENS MED CENT.

Kunkel LM, Monaco A, Hoffman EP, Koenig M;

WPI; 1989-220587/30.

P-PSDB; AAP90373.

Muscular dystrophy gene - used for prepn. of probes, dystrophin  
 polypeptide and antibodies for diagnosis and therapy of muscular  
 dystrophy

Disclosure; Fig 5; 68pp; English.

The inventors claim an MD probe comprising a purified ss NA-SQ which  
 hybridises to at least a part of the MD gene; pure dystrophin (DS)  
 polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The  
 CC probes are equal to or greater than 10b of one of 12 cDNA sequences  
 CC deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd  
 CC gene.  
 CC  
 XX  
 SQ Sequence 12923 BP: 4296 A; 2613 C; 2963 G; 3044 T; 7 other;

Query Match 82.0%; Score 16.4; DB 10; Length 12923;  
 Best Local Similarity 94.4%; Pred. No. 1.5e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ccagttctcaagagcagc 19

DB 11904 CAGTCTCAATGACGAC 11887  
 ||||| ||||| ||||| |||||

RESULT 9  
 AA125968/C  
 ID AA125968 standard; DNA; 209 BP.  
 XX  
 AC AA125968;

12-OCT-2001 (first entry)

Probe #15901 for gene expression analysis in human cervical cell sample.

Probe: human; microarray; gene expression; cervical epithelial cell;  
 cervical cancer; ss.

Homo sapiens.

W0200157278-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00670.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR.

WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -

Claim 25; SEQ ID NO 15901; 487pp; English.

The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging  
 CC of diseases of the cervix, notably cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 209 BP: 47 A; 59 C; 54 G; 49 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 209;  
 Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 19  
 ||| ||||| ||||| |||||  
 DB 175 CCGGCTCTCAAGAGCAGC 157

RESULT 10

AA153359/C  
 ID AA153359 standard; DNA; 209 BP.

AA153359;

DF 17-OCT-2001 (first entry)  
XX DE Probe #22045 used to measure gene expression in human placenta sample.  
XX KW Probe: microarray: human; placenta; antenatal diagnosis:  
XX genetic disorder; ss.  
OS Homo sapiens.  
XX MO200157272-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00663.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488897/53.  
XX DR Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human placenta -  
XX PS Claim 25: SEQ ID No 22045; 654bp; English.  
XX CC The present invention relates to single exon nucleic acid probes (SENP).  
XX CC The present sequence is one such probe. The probes are useful for  
XX CC producing a microarray for predicting, measuring and displaying gene  
XX CC expression in samples derived from human placenta. The probes are useful  
XX CC for antenatal diagnosis of human genetic disorders.  
SQ Sequence 209 BP; 47 A; 59 C; 54 G; 49 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 209;  
Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 ccaattctcaagagcag 19  
175 CCGGCTCTCAAGAGCAG 157

RESULT 11  
AA116773/C  
ID AA116773 standard; DNA: 563 BP.  
XX AA116773;  
XX AC 12-OCT-2001 (first entry)  
XX DT Probe #6706 for gene expression analysis in human cervical cell sample.  
XX DE Probe: human; microarray; gene expression; cervical epithelial cell;  
XX KW cervical cancer; ss.  
XX OS Homo sapiens.  
XX PN MO200157272-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00670.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488901/53.  
XX DR Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human cervical epithelial cells -  
XX PS Claim 25; SEQ ID No 6706; 487bp; English.  
XX CC The present invention relates to human single exon nucleic acid probes  
XX CC (SENP). The present sequence is one such probe. The SENPs are derived  
XX CC from human Hela cells. The SENPs can be used to produce a single exon  
XX CC microarray, which can be used for measuring human gene expression in a  
XX CC sample derived from human cervical epithelial cells. By measuring gene  
XX CC expression, the probes are therefore useful in grading and/or staging  
XX CC of diseases of the cervix, notably cervical cancer.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 563 BP; 115 A; 153 C; 178 G; 117 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 563;  
Best Local Similarity 89.5%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 ccaattctcaagagcag 19  
326 CCGGCTCTCAAGAGCAG 308

RESULT 12  
AA140280/C  
ID AA140280 standard; DNA: 563 BP.  
XX AA140280;  
XX AC 17-OCT-2001 (first entry)  
XX DT Probe #8966 used to measure gene expression in human placenta sample.  
XX DE Probe: microarray: human; placenta; antenatal diagnosis:  
XX KW genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN MO200157272-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00663.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.



```

OY      1 ccagttctcaagagcag 19
          |||||  |||||  |||||
Db      172 ccagttcacaagtagcag 190

```

RESULT 15

ID	AAH00627	standard; DNA; 673 .BP.
XX		

AC AAH00627;

DT 24-JUL-2001 (first entry)  
 YY

Enterococcus avium nucleotide sequence SEQ ID NO:618

KM Species specific; genus specific; family specific; probe; detection;  
identification; algal; archaeal; bacterial; fungal; parasitoid

KW microorganism; diagnosis; translation elongation factor Tu; toxin,  
translation elongation factor G; BCCA recombinant; restriction

catalytic subunit of proton-translocating ATPase; antimicrobial, vaccine; primer; de

Enterococcus avium

XX  
PN  
W0200123604-A?

XX  
PD 05-APR-2001

XX 28-SEP-2000: 2000WC  
PF

XX  
PR 28-SEP-1999.

PR 19-MAY-2000  
XX

PA (INF-) INFECTION DIAGNOSTIC  
XX XX

PI Bergeron M, Boissinot M,  
PI Picard FJ, Roy PH:

DR WPT: 2001-245006/25.

aa	PT	Nucleic acid sequence
1	1	5'-GAG-3'
2	2	5'-GAG-3'
3	3	5'-GAG-3'
4	4	5'-GAG-3'
5	5	5'-GAG-3'
6	6	5'-GAG-3'
7	7	5'-GAG-3'
8	8	5'-GAG-3'
9	9	5'-GAG-3'
10	10	5'-GAG-3'
11	11	5'-GAG-3'
12	12	5'-GAG-3'
13	13	5'-GAG-3'
14	14	5'-GAG-3'
15	15	5'-GAG-3'
16	16	5'-GAG-3'
17	17	5'-GAG-3'
18	18	5'-GAG-3'
19	19	5'-GAG-3'
20	20	5'-GAG-3'
21	21	5'-GAG-3'
22	22	5'-GAG-3'
23	23	5'-GAG-3'
24	24	5'-GAG-3'
25	25	5'-GAG-3'
26	26	5'-GAG-3'
27	27	5'-GAG-3'
28	28	5'-GAG-3'
29	29	5'-GAG-3'
30	30	5'-GAG-3'
31	31	5'-GAG-3'
32	32	5'-GAG-3'
33	33	5'-GAG-3'
34	34	5'-GAG-3'
35	35	5'-GAG-3'
36	36	5'-GAG-3'
37	37	5'-GAG-3'
38	38	5'-GAG-3'
39	39	5'-GAG-3'
40	40	5'-GAG-3'
41	41	5'-GAG-3'
42	42	5'-GAG-3'
43	43	5'-GAG-3'
44	44	5'-GAG-3'
45	45	5'-GAG-3'
46	46	5'-GAG-3'
47	47	5'-GAG-3'
48	48	5'-GAG-3'
49	49	5'-GAG-3'
50	50	5'-GAG-3'
51	51	5'-GAG-3'
52	52	5'-GAG-3'
53	53	5'-GAG-3'
54	54	5'-GAG-3'
55	55	5'-GAG-3'
56	56	5'-GAG-3'
57	57	5'-GAG-3'
58	58	5'-GAG-3'
59	59	5'-GAG-3'
60	60	5'-GAG-3'
61	61	5'-GAG-3'
62	62	5'-GAG-3'
63	63	5'-GAG-3'
64	64	5'-GAG-3'
65	65	5'-GAG-3'
66	66	5'-GAG-3'
67	67	5'-GAG-3'
68	68	5'-GAG-3'
69	69	5'-GAG-3'
70	70	5'-GAG-3'
71	71	5'-GAG-3'
72	72	5'-GAG-3'
73	73	5'-GAG-3'
74	74	5'-GAG-3'
75	75	5'-GAG-3'
76	76	5'-GAG-3'
77	77	5'-GAG-3'
78	78	5'-GAG-3'
79	79	5'-GAG-3'
80	80	5'-GAG-3'
81	81	5'-GAG-3'
82	82	5'-GAG-3'
83	83	5'-GAG-3'
84	84	5'-GAG-3'
85	85	5'-GAG-3'
86	86	5'-GAG-3'
87	87	5'-GAG-3'
88	88	5'-GAG-3'
89	89	5'-GAG-3'
90	90	5'-GAG-3'
91	91	5'-GAG-3'
92	92	5'-GAG-3'
93	93	5'-GAG-3'
94	94	5'-GAG-3'
95	95	5'-GAG-3'
96	96	5'-GAG-3'
97	97	5'-GAG-3'
98	98	5'-GAG-3'
99	99	5'-GAG-3'
100	100	5'-GAG-3'

archaeal, bacterial, fungal and parasitological species in a test set

PS Claim 24; Page 763;

CC The present invention

and/or primers are derived. The method comprises amplifying the nucleic acid sequence of the gene from which probes

CC species with a combination of defined primer pairs. The method can be

CC related microorganisms e.g. algae, archaea, bacteria, fungi and

CC detection and identification of an algal, archaeal, bacterial, fungal

CC obtained using the method of the invention can be used for the universal

CC detection of at least one antimicrobial agent resistance gene or at

ubiquitous detection and for identification of *Streptococcus pneumoniae*

microorganisms. Microbial species or genus or family or phylum or group

CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,

CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests

CC be determined in an hour and improved accuracy is also achieved.

CC which are given in the exemplification of the present invention.

50 Sequence 673 BP; 210 A; 120 C; 154 G; 189 T; 0 other;

```

Query Match          79.0%; Score 15.8; DB 22; Length 673;
Best Local Similarity 89.5%; Pred. No. 2 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY      1 ccagttctcaaaagagcaq 19
          |||||
Db      205 ccagttctcaaaagttcqa 223

```

Search completed: March 4, 2002, 05:21:23  
Job time: 17512 sec

Mon Mar 4 05:51:52 2002

us-08-956-991-5.rml

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:11:29 : Search time 226.99 Seconds  
(without alignments)  
19.955 Million cell updates/sec

Title: US-08-956-991-5  
Perfect score: 20  
Sequence: 1 ccaggtctcaagaagcagcagg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/Backfillseq1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.2	76.0	3417	2	US-08-464-402-1
2	15.2	76.0	3417	4	US-09-054-775C-1
3	15.2	76.0	4337	3	US-09-187-049-1
4	15.2	76.0	8298	5	PCT-US93-03076-1
5	15.2	76.0	12752	2	US-08-459-146-1
6	15.2	76.0	12752	2	US-08-459-065-1
7	14.8	74.0	656	4	US-09-328-111-313
8	14.8	74.0	1416	4	US-09-234-393-26
9	14.8	74.0	1785	4	US-09-234-393-51
10	14.8	74.0	1785	4	US-09-234-393-53
11	14.8	74.0	1977	4	US-09-234-393-23
12	14.8	74.0	2071	3	US-09-023-023-1
13	14.8	74.0	8535	3	US-08-716-351A-1
14	14.8	74.0	920	3	US-09-258-373-2
15	14.4	72.0	1314	1	US-07-662-005A-1
16	14.4	72.0	1875	3	US-09-258-373-21
17	14.4	72.0	2135	3	US-08-581-148C-17
18	14.4	72.0	2135	1	US-08-053-614-1
19	14.4	72.0	3648	1	US-08-316-397B-1
20	14.4	72.0	3648	2	US-09-034-306-1
21	14.4	72.0	3648	2	US-09-259-437-1
22	14.4	72.0	3648	5	PCT-US93-09782-1
23	14.4	72.0	4821	1	US-08-053-614-3
24	14.4	72.0	4821	1	US-08-316-397B-3
25	14.4	72.0	4821	2	US-09-034-306-3
26	14.4	72.0	4821	3	US-09-259-437-3
27	14.4	72.0	4821	3	US-09-259-437-3

28	14.4	72.0	4821	5	PCT-US93-09782-3	Sequence 3, Appl1
29	14.4	72.0	5925	3	US-08-470-260-4	Sequence 4, Appl1
30	14.4	72.0	5925	3	US-08-471-491-4	Sequence 4, Appl1
31	14.4	72.0	5925	3	US-08-466-662-4	Sequence 1, Appl1
32	14.4	72.0	10299	2	US-08-477-451-1	Sequence 5, Appl1
33	14.4	72.0	10299	2	US-08-477-451-5	Sequence 25, Appl1
34	14.4	72.0	19932	2	US-08-477-451-25	Sequence 40, Appl1
35	14.4	71.0	401	4	US-09-221-298-40	Sequence 19, Appl1
36	14.2	71.0	460	2	US-08-487-727A-1	Sequence 19, Appl1
37	14.2	71.0	580	2	US-08-272-255-19	Sequence 697, App
38	14.2	71.0	580	5	PCT-US93-08565-19	Sequence 410, App
39	14.2	71.0	688	4	US-08-998-416-697	Sequence 2, Appl1
40	14.2	71.0	870	4	US-09-461-697-410	Sequence 1, Appl1
41	14.2	71.0	874	2	US-08-850-119-2	Sequence 402, App
42	14.2	71.0	911	1	US-08-745-603-1	Sequence 396, App
43	14.2	71.0	936	4	US-09-461-697-402	Sequence 396, App
44	14.2	71.0	1017	4	US-09-461-697-396	Sequence 396, App
45	14.2	71.0	1128	4	US-09-461-697-394	Sequence 394, App

#### ALIGNMENTS

RESULT 1  
US-08-464-402-1  
Sequence 1, Application US/08464402  
Patent No. 5838705  
GENERAL INFORMATION:  
APPLICANT: WEI, ET AL.  
TITLE OF INVENTION: Human DNA Ligase III  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,402  
FILING DATE: June 5, 1995  
CLASSIFICATION: 4.35  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01939  
FILING DATE: 41 MAR 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-388  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1740  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3417 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-08-464-402-1

Query Match 76.0% Score 15.2; DB 2; Length 3417;  
Best Local Similarity 85.0% Pred. No. 71;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 2559 CCACTTGTCCAGAGAGG 2578

## RESULT 2

US-09-054-775C-1  
Sequence 1, Application US/09054775C  
Patent No. 6284504  
GENERAL INFORMATION:  
APPLICANT: Wei, Ying-Fei  
Yu, Guo-Liang  
Haseltine, William  
TITLE OF INVENTION: Human DNA Ligase III  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/054,775C  
FILING DATE: 03-Apr-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/464,402  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: PCT/US95/03939  
FILING DATE: 31-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Kenley K.  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: P1161D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3417 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-054-775C-1

Query Match 76.0%; Score 15.2; DB 4; Length 3417;  
Best Local Similarity 85.0%; Pred. No. 71;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccagltctcaagagcagg 20  
||||| ||||| |||  
Db 2559 CCACTTGTCCAGAGAGG 2578

RESULT 3  
US-09-187-049-1  
Sequence 1, Application US/09187049  
Patent No. 6117666  
GENERAL INFORMATION:  
APPLICANT: Lampka, Gayle K.  
TITLE OF INVENTION: PLASTID PROTEOLYTIC PROCESSING ENZYME  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BRINKS HOFER GILSON & LIONE

STREET: P.O. Box 10395  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/187,049  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/695,177  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Martin, Alice O.  
REGISTRATION NUMBER: 35,601  
REFERENCE/DOCKET NUMBER: 7814/16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312 321-4200  
TELEFAX: 312 321-4299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
US-09-187-049-1

Query Match 76.0%; Score 15.2; DB 3; Length 4337;  
Best Local Similarity 85.0%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccagltctcaagagcagg 20  
||||| ||||| |||||  
Db 2567 CCACTTGTCCAGAGAGG 2586

RESULT 4  
PCT-US93-03076-1  
Sequence 1, Application PC/TUS9303076  
GENERAL INFORMATION:  
APPLICANT: Whitehead Institute for Biomedical Research  
TITLE OF INVENTION: GAV-Associated protein p190 and  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: 2 Millia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/03076  
FILING DATE: 19930331  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WHI92-03A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8298 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 731..5272  
PCT-US93-03076-1

Query Match 76.0%; Score 15.2; DB 5; Length 8298;  
Best Local Similarity 85.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcagc 20  
11111111111111111111  
3763 CCAAGTTCACATGAGCAGC 3782

## RESULT 5

US-08-459-146-1  
Sequence 1, Application US/08459146  
Patent No. 5868405  
GENERAL INFORMATION:  
APPLICANT: Choi, Gil Ho  
APPLICANT: Nuss, Donald Lee  
TITLE OF INVENTION: Genetically Engineered Transmissible  
TITLE OF INVENTION: Hypovirulence  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,146  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/832,117  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R  
REGISTRATION NUMBER: 34,240  
REFERENCE/DOCKET NUMBER: 8589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12752 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Endothia parasitica (Cryptomycetaria)  
ORGANISM: parasitica  
STRAIN: EP713  
US-08-459-146-1

Query Match 76.0%; Score 15.2; DB 2; Length 12752;  
Best Local Similarity 85.0%; Pred. No. 88;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcagc 20  
11111111111111111111  
Db 12192 CCAAGTTCACATGAGCAGC 12211

## RESULT 6

US-08-459-065-1  
Sequence 1, Application US/08459065  
Patent No. 5882642  
GENERAL INFORMATION:  
APPLICANT: Choi, Gil Ho  
APPLICANT: Nuss, Donald Lee  
TITLE OF INVENTION: Genetically Engineered Transmissible  
TITLE OF INVENTION: Hypovirulence  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,065  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/832,117  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R  
REGISTRATION NUMBER: 34,240  
REFERENCE/DOCKET NUMBER: 8589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12752 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Endothia parasitica (Cryptomycetaria)  
ORGANISM: parasitica  
STRAIN: EP713  
US-08-459-065-1

Query Match 76.0%; Score 15.2; DB 2; Length 12752;  
Best Local Similarity 85.0%; Pred. No. 88;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcagc 20  
11111111111111111111  
Db 12192 CCAAGTTCACATGAGCAGC 12211

## RESULT 7

```
US-09-328-111-313
: Sequence 313, Application US/09328111
: Patent No. 6262333
: GENERAL INFORMATION:
: APPLICANT: Endege, Wilson O.
: APPLICANT: Steinmann, Kathleen E.
: APPLICANT: Astle, Jon H.
: APPLICANT: Burgess, Christopher C.
: APPLICANT: Bushnell, Steven E.
: APPLICANT: Carroll III, Eddie
: APPLICANT: Catino, Theodore J.
: APPLICANT: Dertl, Adnan
: APPLICANT: Ford, Donna M.
: APPLICANT: Lewis, Marcia E.
: APPLICANT: Monahan, John E.
: APPLICANT: Schlegel, Robert
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: FILE REFERENCE: CCD-257 (US)
: CURRENT APPLICATION NUMBER: US/09/328.111
: EARLIER FILING DATE: 1998-06-10
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 313
: LENGTH: 656
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(656)
: OTHER INFORMATION: n = A,T,C or G
US-09-328-111-313

Query Match          74.0%; Score 14.8; DB 4; Length 656;
Best Local Similarity 88.9%; Pred. No. 85;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 agtctcaagaagcagc 20
||||| ||||| ||
Db 219 agtctcaagaagcgcg 236

RESULT 8
: Sequence 26, Application US/09234393A
: Patent No. 6265639
: GENERAL INFORMATION:
: APPLICANT: Croteau, Rodney B
: APPLICANT: Crock, John E
: APPLICANT: Bohlman, Jorg
: APPLICANT: Jettler, Reinhard
: APPLICANT: Steele, Christopher L
: TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
: FILE REFERENCE: WSUR113345
: CURRENT APPLICATION NUMBER: US/09/234.393A
: EARLIER FILING DATE: 1999-01-20
: EARLIER APPLICATION NUMBER: 60/072.204
: EARLIER FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 26
: LENGTH: 1416
: TYPE: DNA
: ORGANISM: Abies grandis
US-09-234-393-26

Query Match          74.0%; Score 14.8; DB 4; Length 1416;
Best Local Similarity 88.9%; Pred. No. 97;

US-09-328-111-313
: Sequence 313, Application US/09328111
: Patent No. 6262333
: GENERAL INFORMATION:
: APPLICANT: Endege, Wilson O.
: APPLICANT: Steinmann, Kathleen E.
: APPLICANT: Astle, Jon H.
: APPLICANT: Burgess, Christopher C.
: APPLICANT: Bushnell, Steven E.
: APPLICANT: Carroll III, Eddie
: APPLICANT: Catino, Theodore J.
: APPLICANT: Dertl, Adnan
: APPLICANT: Ford, Donna M.
: APPLICANT: Lewis, Marcia E.
: APPLICANT: Monahan, John E.
: APPLICANT: Schlegel, Robert
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: FILE REFERENCE: CCD-257 (US)
: CURRENT APPLICATION NUMBER: US/09/328.111
: EARLIER FILING DATE: 1998-06-10
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 313
: LENGTH: 656
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(656)
: OTHER INFORMATION: n = A,T,C or G
US-09-328-111-313

Query Match          74.0%; Score 14.8; DB 4; Length 656;
Best Local Similarity 88.9%; Pred. No. 85;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 agtctcaagaagcagc 20
||||| ||||| ||
Db 219 agtctcaagaagcgcg 236

RESULT 9
: Sequence 49, Application US/09234393A
: Patent No. 6265639
: GENERAL INFORMATION:
: APPLICANT: Croteau, Rodney B
: APPLICANT: Crock, John E
: APPLICANT: Bohlman, Jorg
: APPLICANT: Jettler, Reinhard
: APPLICANT: Steele, Christopher L
: TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
: FILE REFERENCE: WSUR113345
: CURRENT APPLICATION NUMBER: US/09/234.393A
: EARLIER FILING DATE: 1999-01-20
: EARLIER APPLICATION NUMBER: 60/072.204
: EARLIER FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 49
: LENGTH: 1785
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:
: OTHER INFORMATION: computer-generated nucleic acid sequence encoding
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (4)..(1782)
US-09-234-393-49

Query Match          74.0%; Score 14.8; DB 4; Length 1785;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 agtctcaagaagcagc 20
||||| ||||| ||
Db 1616 agtctcaagaagcagc 1633

RESULT 10
: Sequence 51, Application US/09234393A
: Patent No. 6265639
: GENERAL INFORMATION:
: APPLICANT: Croteau, Rodney B
: APPLICANT: Crock, John E
: APPLICANT: Bohlman, Jorg
: APPLICANT: Jettler, Reinhard
: APPLICANT: Steele, Christopher L
: TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
: FILE REFERENCE: WSUR113345
: CURRENT APPLICATION NUMBER: US/09/234.393A
: EARLIER FILING DATE: 1999-01-20
: EARLIER APPLICATION NUMBER: 60/072.204
: EARLIER FILING DATE: 1998-01-22
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 51
: LENGTH: 1785
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
```



Query Match	74.0%;	Score 14.8;	DB 3;	Length 2071;
Best Local Similarity	88.9%;	Pred. No. 1e+02;		
Matches 16;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0;

OY 3 agtctcaagagcag 20  
 ||| |||||  
 DB 1545 AGTGATCAAGCAGCAG 1528

RESULT 14

US-08-716-351A-1  
 : Sequence 1, Application US/08716351A  
 : Patent No. 6033905  
 : GENERAL INFORMATION:

APPLICANT:  
 TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based  
 TITLE OF INVENTION: Retroviral Vectors  
 NUMBER OF SEQUENCES: 5  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/716,351A  
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/03784  
 FILING DATE: 06-APR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 15280-128-1PC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 543-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8535 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: 1..8535  
 OTHER INFORMATION: /standard\_name="Galv SEATO Genome"  
 US-08-716-351A-1

Query Match 74.0%; Score 14.8; DB 3; Length 8535;  
 : Local Similarity 88.9%; Pred. No. 1.3e+02;  
 Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 cagtctcaagagcag 19  
 ||| |||||  
 DB 1977 CAGGTCTCAAGJCAGCAG 1994

RESULT 15

US-09-258-373-2/c  
 : Sequence 2, Application US/09258373  
 : Patent No. 6150110  
 : GENERAL INFORMATION:

APPLICANT: Fletcher, Jonathan A.  
 TITLE OF INVENTION: HMG(Y)-LAMA4\* FUSION ONCOGENE,  
 TITLE OF INVENTION: ONCOPROTEIN AND METHODS OF USE  
 FILE REFERENCE: B0801/7135/ERP  
 CURRENT APPLICATION NUMBER: US/09/258,373  
 CURRENT FILING DATE: 1999-02-26  
 EARLIER APPLICATION NUMBER: 60/076,401  
 EARLIER FILING DATE: 1998-02-28  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 2  
 : LENGTH: 920  
 : TYPE: DNA  
 : ORGANISM: Homo Sapiens  
 US-09-258-373-2

Query Match 72.0%; Score 14.4; DB 3; Length 920;  
 Best Local Similarity 93.8%; Pred. No. 1.4e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ttctcaagagcag 20  
 |||||  
 DB 225 TTCTCTAAGCAGCAG 210

Search completed: March 4, 2002, 05:11:33  
 Job time: 16947 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:35 : Search time 6206.58 Seconds  
(without alignments)  
34.627 Million cell updates/sec

Title: US-08-956-991-6

Perfect score: 20

Sequence: 1 cctgatcactgcaggaag 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1:51937 segs, 537289281 residues 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_estnum:\*  
3: em\_estlin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estda:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hnc:\*  
10: 9b\_estl:\*  
11: 9b\_est2:\*  
12: 9b\_hnc:\*  
13: 9b\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_hiv:\*  
17: em\_gss\_pla:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	744	11	BF348620 602019924
2	18.4	92.0	424	13	AZ297511 KFCI-23-1
3	17.4	87.0	903	13	CNS05070
4	17	85.0	559	11	BG700618
5	17	85.0	655	11	BG701055
6	16.8	84.0	176	10	BR373590
7	16.8	84.0	188	11	F25472
8	16.8	84.0	270	10	BB528026
9	16.8	84.0	274	10	BB367406
10	16.8	84.0	276	10	BE032419
11	16.8	84.0	278	10	BB604281
12	16.8	84.0	291	10	BB429902

13	16.8	84.0	292	11	BF825905	BF825905 CM4-HN002
14	16.8	84.0	298	10	BB116563	BB116563
15	16.8	84.0	334	10	AI841727	AI841727 UI-M-ALO-
16	16.8	84.0	338	11	F00929	F00929 HSR826032.S
17	16.8	84.0	339	10	AA855923	AA855923 vW80e12.r
18	16.8	84.0	349	10	AI836340	AI836340 UI-M-A00-
19	16.8	84.0	349	11	R05334	AI905927 PM-BT103-
20	16.8	84.0	350	10	AI905927	AI905927 PM-BT103-
21	16.8	84.0	358	11	BF61521	BF61521 274319 MA
22	16.8	84.0	359	11	BF825966	BF825966 CM4-HN002
23	16.8	84.0	373	11	BE82657	BE82657 156271 MA
24	16.8	84.0	381	10	AV592640	AV592640 AV592640
25	16.8	84.0	388	10	AJ272002	AJ272002 AJ272002
26	16.8	84.0	401	10	AJ278473	AJ278473 AJ278473
27	16.8	84.0	401	11	BF728908	BF728908 PM2-KT000
28	16.8	84.0	416	10	AA65681	AA65681 VP95B03.r
29	16.8	84.0	428	11	BF826523	BF826523 CM4-HN002
30	16.8	84.0	445	11	BF795797	BF795797 U7SM-SM37
31	16.8	84.0	448	11	BF873693	BF873693 IL5-ET011
32	16.8	84.0	450	11	BF826402	BF826402 CM4-HN002
33	16.8	84.0	450	11	BE610562	BE610562 326654 MA
34	16.8	84.0	459	10	BE233426	BE233426 139605 MA
35	16.8	84.0	460	10	AJ278546	AJ278546 AJ278546
36	16.8	84.0	466	10	AW535840	AW535840 I0232D11-
37	16.8	84.0	467	13	AZ851563	AZ851563 2M0153B22
38	16.8	84.0	468	11	BF826596	BF826596 CM4-HN002
39	16.8	84.0	471	11	BG796032	BG796032 U7SM-SM39
40	16.8	84.0	478	11	BF592276	BF592276 7153607.x
41	16.8	84.0	485	11	BF599809	BF599809 264128 MA
42	16.8	84.0	485	11	AOB15436	AOB15436 HS-5291.B
43	16.8	84.0	487	11	BF603509	BF603509 269044 MA
44	16.8	84.0	490	11	BG794127	BG794127 U7SM-SM19
45	16.8	84.0	490	11	BG794127	BG794127

#### ALIGNMENTS

RESULT 1  
LOCUS BF348620 744 bp mRNA  
DEFINITION 602019924F1 NCI CGAP\_Brn67 Homo sapiens CDNA clone IMAGE:4155508  
5', mRNA sequence.  
ACCESSION BF348620  
VERSION BF348620.1 GI:11296215  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-femail.nih.gov  
Tissue: Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM9426 row: p column: 05  
High quality sequence stop: 697.  
Location/Qualifiers  
1. 744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_1lb="NCI-CGAP\_Brn67"  
/rlnuc\_type="amplastic oligodendrogloma with 1p/19q  
loss"

#### FEATURES

source

```

BASE COUNT      188 a      196 c      210 g      150 t
ORIGIN
    /lab.host="DH10B (T1 phage-resistant)"
    /note="Jorgan; brain; Vector: PCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
    Average insert size 2.3 kb. Constructed by Life
    Technologies. Note: this is a NCI-CGAP Library."

```

Query Match	100.0%	Score 20;	DB 11;	Length 744;
Best Local Similarity	100.0%	Pred. No. 72;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

QY      1 cclgtatgacctgcaygaag 20
        |||||
Db      195 CCTGTATGACCTGCAGGAAG 214

```

RESULT	2
A <sub>i</sub>	511
DEFINITION	AC297511 424 bp DNA GSS 27-JUL-2000 RPCR-23-117G10.TJ RPCR-23 Mus musculus genomic clone RPCR-23-117G10 , DNA s'quence.

ACCESSION	AZ297511	GI:95392296
VERSION	AZ297511.1	
KEYWORDS	GSS.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

**ORGANISM** Mus musculus; Mus musculus; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
**REFERENCE** 1 (bases 1 to 424)  
**AUTHORS** Zhao, S., Niernman, W., Feldblum, T., Malek, J., Shatsman, S., Alintet, A., Krawinkel, B., Tevase, G., Geer, Jr., Kroil, M., de Jong,

TITLE	Mouse BAC End Sequences from Library RPCI-23
JOURNAL	Unpublished (1999)
COMMENT	Other_GSSs: RPCI-23-117G10.TV

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel.: 301 838 0200  
Fax: 301 838 0208  
Email: [szhaoc@tigr.org](mailto:szhaoc@tigr.org)  
Clones are derived from the mouse BAC library RPc1-23. For BAC library availability, please contact pjefer de jong ([pjefer@dejong.med.buffalo.edu](mailto:pjefer@dejong.med.buffalo.edu)). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics ([info@resgen.com](http://resgen.com)). BAC end page: [http://www.tigr.org/tdb/bac-ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac-ends/mouse/bac_end_intro.html)  
plate: 117 row: G column: 10  
Seq primer: SP6  
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .424

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPC1-23-117G10"
/clone_lib="RPC1-23"
/sex="Female"
/lab_host="DH10B"
/notes="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1;
ECORI: Site: 2; ECORI: Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of ECORI and ECORI Methylase. Site
selected DNA was cloned into the pBAC3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL, Life Technologies)."
BASE COUNT
101 a 99 c 130 g 94 t
ORIGIN

```

Query Match	Score	DB	Length
Best Local Similarity	92.0%	13	424
Best Local Similarity	95.0%	Pred. No. 3.5e+02	
Matches	19	Conservative	0
		Mismatches	1
		Indels	0
		Gaps	0

RESULT	3
CNS05070	
LOCUS	903 bp DNA
DEFINITION	CNS05070 26-JUL-2000 Tetradion nigriviridis genome survey sequence T7 of clone 027C17 of library B from Tetradion nigriviridis, genomic survey

ACCESSION AL315219  
VERSION AL315219.1 GI:9548107  
KEYWORDS GSS: genome survey sequence.  
SOURCE Tetradon nigroviridis.  
ORANISM Tetradon nigroviridis.  
Chordata: Vertebrata: Euteleostomi:

**REFERENCE**  
**AUTHORS**

1 (bases 1 to 903)  
Roest Croillius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Rottler, P., Queller, F.

**TITLE** Estimate of human gene number provided by genome-wide analysis of Tetraodon nigroviridis DNA sequence  
**JOURNAL** Nat. Genet. 25 (2), 235-238 (2000)

REFERENCE	AUTHORS	TITLE	JOURNAL
2 (bases 1 to 903)	Crollius, H.R., Jallion, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C.		
	Fischer, C., Bonneau, I., Billault, A., Queller, F., Saurin, W., Bernot, A. and Weissbach, J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish tetraodon nigroviridis	Genome Res. 10 (7), 939-949 (2000)

**TITLE** Direct Submission  
**JOURNAL** Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases  
**COMMENT** This sequence is a single read and was generated as part of a large  
set of the metagenomic microviridis

genome. For more information, please take a look at <http://www.genoscope.cns.fr/16eladon>.

FEATURES  
source  
1. .903  
local ion/qualifiers

```

/origansim="telodon trigloporus"
/db_xref="taxon:99883"
/clone="027c17"
/clone_1lt="B"
/note="genoscope sequence ID : COAB027AB09C1-end : T7"

```

BASE COUNT ORIGIN	302 a	159 c	176 g	231 l	330 m
1	100	100	100	100	100
2	100	100	100	100	100
3	100	100	100	100	100
4	100	100	100	100	100
5	100	100	100	100	100
6	100	100	100	100	100
7	100	100	100	100	100
8	100	100	100	100	100
9	100	100	100	100	100
10	100	100	100	100	100
11	100	100	100	100	100
12	100	100	100	100	100
13	100	100	100	100	100
14	100	100	100	100	100
15	100	100	100	100	100
16	100	100	100	100	100
17	100	100	100	100	100
18	100	100	100	100	100
19	100	100	100	100	100
20	100	100	100	100	100
21	100	100	100	100	100
22	100	100	100	100	100
23	100	100	100	100	100
24	100	100	100	100	100
25	100	100	100	100	100
26	100	100	100	100	100
27	100	100	100	100	100
28	100	100	100	100	100
29	100	100	100	100	100
30	100	100	100	100	100
31	100	100	100	100	100
32	100	100	100	100	100
33	100	100	100	100	100
34	100	100	100	100	100
35	100	100	100	100	100
36	100	100	100	100	100
37	100	100	100	100	100
38	100	100	100	100	100
39	100	100	100	100	100
40	100	100	100	100	100
41	100	100	100	100	100
42	100	100	100	100	100
43	100	100	100	100	100
44	100	100	100	100	100
45	100	100	100	100	100
46	100	100	100	100	100
47	100	100	100	100	100
48	100	100	100	100	100
49	100	100	100	100	100
50	100	100	100	100	100
51	100	100	100	100	100
52	100	100	100	100	100
53	100	100	100	100	100
54	100	100	100	100	100
55	100	100	100	100	100
56	100	100	100	100	100
57	100	100	100	100	100
58	100	100	100	100	100
59	100	100	100	100	100
60	100	100	100	100	100
61	100	100	100	100	100
62	100	100	100	100	100
63	100	100	100	100	100
64	100	100	100	100	100
65	100	100	100	100	100
66	100	100	100	100	100
67	100	100	100	100	100

Query Match	87.0%	Score 17.4	DB 13	Length 903
Best Local Similarity	94.7%	Pred. No. 9.5e+02		
Matches 18	Conservative 0	Mismatches 1	Indels 0	Caps 0

QY	1	cctgtatgaccctgcaggaa	19
Db	69	cctgaatgcacctgcaggaa	87

LOCUS	559 bp	mRNA	EST	07-MAY-2001
DEFINITION	602682215F1 NIH_MGC_95	Homo sapiens	CDNA clone IMAGE:4815146	5'
mRNA sequence.				

ACCESSION BG700618  
 VERSION BG700618.1 GI:13970140  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE: NIH-MGC <http://mgi.nci.nih.gov/>,  
 (bases 1 to 559)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shitaki Toshitsuki and Piero Carninci (RIKEN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10711 row: m column: 03  
 High quality sequence stop: 559.  
 Location/Qualifiers  
 1..559  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH-MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="organ: brain; Vector: pBluescriptK (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to KOF 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."  
 HASH COUNT 175 a 114 c 139 g 131 t  
 ORIGIN  
 Query Match 85.0%; Score 17; DB 11; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 cctgatacctgcag 17  
 |||  
 121 CCTGTATACCTGCAG 105  
 RESULT 5  
 BG701055 655 bp mRNA EST 07-MAY-2001  
 602662072p1 NIH-MGC\_95 Homo sapiens cDNA clone IMAGE:4814849 5'  
 mRNA sequence.  
 ACCESSION BG701055  
 VERSION BG701055.1 GI:13971014  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE: NIH-MGC <http://mgi.nci.nih.gov/>,  
 (bases 1 to 655)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shitaki

Toshitsuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM10710 row: p column: 18  
 High quality sequence stop: 620.  
 Location/Qualifiers  
 1..655  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH-MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="organ: brain; Vector: pBluescriptK (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to KOF 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."  
 BASE COUNT 205 a 140 c 154 g 156 t  
 ORIGIN  
 Query Match 85.0%; Score 17; DB 11; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 cctgatacctgcag 17  
 |||  
 121 CCTGTATACCTGCAG 105  
 RESULT 6  
 BB373590  
 LOCUS BB373590  
 DEFINITION BB373590 176 bp mRNA EST 13-JUL-2000  
 BB373590 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone G130071C07 3' similar to AB025261 Sus scrofa mRNA for myosin heavy chain 2b, mRNA sequence.  
 ACCESSION BB373590  
 VERSION BB373590.1 GI:708084  
 KEYWORDS EST  
 SOURCE mouse  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE: Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M., Matsuyama, T., Miki, K., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yananaka, I., Yano, K., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al.)  
 Unpublished (2000)  
 TITLE Yoshihide Hayashizaki  
 JOURNAL Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216



The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,  
URL: http://genome.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki  
N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Thermolabile and its application for the synthesis of full length  
cDNA. Proc Natl Acad Sci U S A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki  
N., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.riken.go.jp) for  
further details.

#### Location/Qualifiers

```

JRES
source
1. 270
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="D93004H02"
/clone_1lb="RIKEN full-length enriched, 15 days embryo
head"
/sex="mixed"
/tissue-type="head"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
transcribed by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I"
```

```

COUNT 58 a 101 c 55 g 56 t
Query Match 84.0%; Score 16.8; DB 10; Length 270;
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 cctgatagacctgcaggaag 20
||||| ||||| ||||| |||||
db 72 CCTGTCTGACCTCGACGATG 91
```

```

RESULT 9
BR367406 274 bp mRNA EST 12-JUL-2000
LOCUS BR367406 RIKEN full-length enriched, 16 days embryo head Mus
DEFINITION musculus cDNA clone C130037M06 3' similar to AB025261 Sus scrofa
RNA for myosin heavy chain 2b, mRNA sequence.
BR367406
ACCESSION BR367406.1 GI:9079235
VERSION
KEYWORDS house mouse.
SOURCE
ORGANISM Mus musculus
```

#### REFERENCE

1 (bases 1 to 274)  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata  
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, T., Sugahara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya  
T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamane, A.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
M., Muramatsu, M., and Hayashizaki, Y.  
RIKEN Mouse ESTs (Konno, H., et al.)  
Unpublished (2000)

#### JOURNAL

#### COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki  
N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Thermolabile and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki  
N., and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P., and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome.riken.go.jp) for  
further details.

#### FEATURES

##### source

```

1. 274
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C130037M06"
/clone_1lb="RIKEN full-length enriched, 16 days embryo
head"
/sex="mixed"
/tissue-type="head"
/dev_stage="16 days embryo"
/lab_host="DH10B"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
transcribed by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I"
```

#### BASE COUNT

```
87 a 83 c 70 g 34 t
```

#### ORIGIN

ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus. 1 (bases 1 to 278)
REFERENCE	Aizawa, K., Akhita, S., Akinura, T., Arai, A., Arakawa, T., Carninci, P., Iwazawa, K., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodojima, Y., Imocani, K., Ishii, Y., Itoh, M., Izawa, M., Kawaji, J., Kohama, Y., Konno, H., Kusabebe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Salo, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toyota, T., Watanabe, A., Yamamasa, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M., and Hayashizaki, Y.
JOURNAL	RIKEN Mouse ESTs (Aizawa, K. et al. 2000) Unpublished (2000)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9232 Fax: 81-45-503-9216 Email: genome-resgsc.riken.go.jp, URL:http://genome.qsc.riken.go.jp/ Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Thermotranscription and thermostabilization of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itoh, M., Kitsumaru, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y. Accumulated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site ( <a href="http://genome.ritc.riken.go.jp">http://genome.ritc.riken.go.jp</a> ) for further details.
FEATURES	location/qualifiers 1..278 /organism "Mus musculus" /strain "57R1/bal" /db_xref "taxon:10090" /clone "p14015r15" /comment "p14015r15 length enriched, 15 days embryo head" /sex="mixed" /tissue_type="head" /dev_stage="15 days embryo" /note="Site_1: SalI; Site_2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGACAGACAGATTCGACAGCTGTGGTTTTTTTGGTTTAA 3']. cDNA was prepared by using Trehalose thermoactivated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' CGACGAGCATTCGACGATTGAATTAATCCGCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1."
BASE COUNT	79 a 53 c 98 g 48 t
ORIGIN	
Query Match	84.0%; Score 16.8; DB 10; Length 278;
Best Local Similarity	90.0%; Pred. No. 1.7e+03;



Matches 18: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

OY 1 cctgtatgacctgcagaag 20  
||||| ||||| ||||| |||||  
Db 155 CCTGATGACCTCCAGAG 136

RESULT 12  
B8429902 291 bp mRNA EST 18-JUL-2000  
LOCUS B8429902 RIKEN full-length enriched, adult male hippocampus Mus  
DEFINITION musculus cDNA clone C630030M10 3' similar to M28383 Mouse AE3 mRNA,  
mRNA sequence.  
B8429902  
B8429902.1 GI:9269629

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
FLORES

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 291)  
Komuro, H., Aizawa, K., Akahira, S., Akiyama, J., Arahawa, T., Carninci,  
P., Edo, T., Fukuda, S., Fukunishi, H., Hata, A., Hayatsu, N.,  
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,  
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suganaka, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toyota,  
T., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yamanaka, J.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Komuro, H., et al.)  
Unpublished (2000)

TITLE  
JOURNAL  
COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp/  
URL: http://genome-gsc.riken.go.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki,  
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermostabilization and thermocatalysis of the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (http://genome-rtc.riken.go.jp) for  
further details.

FEATURES  
SOURCE

location/Qualifiers  
1..291  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="C630030M10"  
/clone\_1db="RIKEN full-length enriched, adult male  
hippocampus"  
/sex="male"  
/tissue\_type="hippocampus"  
/dev\_stage="adult"  
/lab\_host="SOLR"  
/note="Site 1: XhoI; Site 2: HinfI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia

BASE COUNT  
ORIGIN  
54 a 117 c 51 g 69 t

Query Match 84.0%; Score 16.8; DB 10; Length 291;  
Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
Matches 18: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

OY 1 cctgtatgacctgcagaag 20  
||||| ||||| ||||| |||||  
Db 90 CCTGATGACCTCCAGAG 109

RESULT 13  
B825905 242 bp mRNA EST 13-JAN-2001  
LOCUS B825905  
DEFINITION CM4-HN0020-181100-444-404 HN0020 Homo sapiens cDNA, mRNA sequence.  
ACCESSION B825905  
VERSION B825905.1 GI:12168691  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 242)  
Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R.,  
Nagai, M.A., de Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, R., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
2020/6/63  
Contact: Simpson A.J.V.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Frei Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=CM4&t2=CM4-HN0020-  
181100-444-004&t3=2000-11-18&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 3  
High quality sequence stop: 292.

FEATURES  
SOURCE

location/Qualifiers  
1..292  
/organism="Homo sapiens"  
/strain="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1db="HN0020"  
/dev\_stage="Adult"  
/note="Organ: head, normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORFESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under



```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-AL0-abo-f-04-0-UI"
/clone.lib="NIH_BMAP_MCO"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p7773D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_MCO library is a non-normalized library
constructed from mouse cortex. The tag is a string of 5
nucleotides present between the Not I site and the
oligo-dr track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories.
TAG_LIB=NIH_BMAP_MCO
TAG_TISSUE=prefrontal-cortex
TAG_SEQ=GCACA"
BASE COUNT      60 a      92 c      100 g      82 t
IN

```

```

Query Match      84.0%; Score 16.8; Dh 10; Length 334.
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Caps 0;
Oy      1 cctgtagacctgcaggaag 20
      ||||| ||||| ||||| |||||
Db      215 CCTGTGTGACCTGCAGGATG 196

```

Search completed: March 4, 2002, 02:14:39  
 Job time: 6488 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

UM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:21:23 : Search time 578.64 seconds  
(without alignments)  
29.632 Million cell updates/sec

Title: US-08-956-991-6

Perfect score: 20

Sequence: 1 ccgtatgacctgcaggaag 20

Scoring table:

IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

1 number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N\_Geneseq\_1101.\*  
1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	19	AAV31984 Human Down syndrome
2	20	100.0	388	19	AAV31982 Human Down syndrome
3	20	100.0	6604	19	AAV31981 Human Down syndrome
4	18.4	92.0	6413	19	AAV31988 Human Down syndrome
5	17.4	87.0	2173	19	AAV31987 Human Down syndrome
6	16.8	84.0	1438	21	AAAI1722 Human Down syndrome
7	16.8	84.0	1579	21	AAAI1722 Human Down syndrome
8	16.8	84.0	1954	22	AAAI1722 Human Down syndrome
9	16.8	84.0	1954	22	AAAI1722 Human Down syndrome
10	16.8	84.0	1954	22	AAAI1722 Human Down syndrome
11	16.8	84.0	1954	22	AAAI1722 Human Down syndrome

12	16.8	84.0	1955	22	AAI16328	Probe #5014 used t
13	16.8	84.0	1955	22	AAI16328	Probe #4740 used t
14	16.8	84.0	6016	22	AAH57390	Human skeletal mus
15	16.8	84.0	8948	20	AAZ24616	Human lung cancer
16	16.8	84.0	8948	21	AAZ24616	Human lung cancer
17	16.4	82.0	847	19	AAV59703	Human secreted pro
18	16.4	82.0	852	19	AAV59704	Human secreted pro
19	16.4	82.0	873	21	AAFI1961	Trichoderma reesei
20	16.4	82.0	1250	20	AAAX24827	Human TKCA encodin
21	16.4	82.0	1284	20	AAAX83631	Human htk1 coding
22	16.4	82.0	1287	19	AAV59705	Human secreted pro
23	16.4	82.0	1354	19	AAV59705	Human secreted pro
24	16.4	82.0	1486	19	AAV59705	Human secreted pro
25	16.4	82.0	1624	20	AAAX24828	Calcium activated
26	16.4	82.0	1877	21	AAZ51619	Human membrane cha
27	16.4	82.0	2248	20	AAAX24826	Calcium activated
28	16.4	82.0	2261	20	AAAX24825	Human calcium acti
29	16.4	82.0	2261	21	AAAC36224	Arabidopsis thalia
30	16.4	82.0	2261	21	AAAC36224	Arabidopsis thalia
31	16.4	82.0	2261	21	AAAC36224	Arabidopsis thalia
32	15.8	79.0	263	21	AAAC21515	Human colon cancer
33	15.8	79.0	480	22	AAI110365	Human MLP DNA
34	15.8	79.0	480	22	AAI131618	Human MLP DNA
35	15.8	79.0	480	22	AAI100300	Human MLP DNA
36	15.8	79.0	617	22	AAH36749	Human MLP DNA
37	15.8	79.0	1273	22	AAH48623	Human MLP DNA
38	15.8	79.0	1273	22	AAH48623	Human MLP DNA
39	15.8	79.0	1273	22	AAH48623	Human MLP DNA
40	15.8	79.0	1285	13	AAO20228	Human MLP DNA
41	15.8	79.0	1296	13	AAO20228	Human MLP DNA
42	15.8	79.0	1296	16	AAO80082	Sequence of the mu
43	15.8	79.0	1296	16	AAO80082	Sequence of the mu
44	15.8	79.0	1322	13	AAQ25979	ADPGP mutant g19c
45	15.8	79.0	1323	19	AAV23650	g19c. Escherichia

#### ALIGNMENTS

RESULT 1	
ID	AAV31984 standard; cDNA: 20 bp.
XX	AAV31984
XX	AAV31984
XX	28-SEP-1998 (first entry)
XX	Human Down syndrome-cell adhesion molecule DS-CAM primer B9-131R.
XX	DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;
XX	signal transduction; trisomy 21; mental retardation;
XX	holoprosencephaly; corpus callosum agenesis;
XX	schizencephaly; diagnosis; assay; human; PCR; primer: SS.
XX	Synthetic.
XX	Homo sapiens.
XX	WO9817795-A1.
XX	30-APR-1998.
XX	23-OCT-1997: 97MO-US19547.
XX	25-OCT-1996: 96US-0029322.
XX	(CEDA-) CEDARS SINAI MEDICAL CENT.
XX	Korenberg JR;
XX	WPI: 1998-271791/24.
XX	New isolated human's Syndrome-cell adhesion molecule - used to
XX	develop products for detection, diagnosis and therapy of
XX	PT

pr developmental and neurological abnormalities  
 xx  
 xx  
 ps  
 cc Example 5; page 79; 109pp: English.  
 xx  
 cc Reverse primer B9-131R and forward primer B9-131F (see AAV31983)  
 cc were used in RT-PCR assays of cDNA libraries of various human  
 cc tissues to determine human Down syndrome-cell adhesion molecule  
 cc DS-CAM (see AAV31981 and AAV31988) expression. The results  
 cc demonstrated expression of human DS-CAM mRNA in foetal and adult  
 cc brain, and foetal kidney. In addition, a breast carcinoma cell  
 cc line showed expression of human DS-CAM mRNA. DS-CAM polypeptides  
 cc (see AAV42066-87) are associated with developmental and neurological  
 cc processes. The polypeptides and nucleic acids are used to develop  
 cc products for the detection, diagnosis and therapy of developmental  
 cc and neurological abnormalities.  
 xx  
 xx  
 SO Sequence 20 BP: 5 A: 5 C: 6 G: 4 T: 0 other:

1	1	100.0%	Score 20;	PI 19;	length 20;
1	1	100.0%	Pred. No. 4 5;		
Matches	20;	Conservative	0;	Mismatches	0;
Indels					0;
Gaps					0;

RESULT	2
AAV31982/c	
ID	AAV31982 standard; cDNA; 388 BP
XX	
AC	AAV31982;
XX	
DT	28-SEP-1998 (first entry)

XX	Human Down syndrome-cell adhesion molecule DS-AM probe c5
DE	
XX	
KM	DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;
KM	signal transduction; trisomy 21; mental retardation;
KM	holoprosencephaly; corpus callosum agenesis;
KM	schizencephaly; diagnosis; assay; human; ds; ss.

...	Homo sapiens.	
OS	XX	
PN	MO9817795-A1.	
XX	P.	
P	30-APR-1998.	
λ		
Ph	23-OCT-1997:	97MO-US19547
XX		
PR	25-OCT-1996:	96US-0029322

AA  
PA  
(CEDA-) CEDARS SINAI MEDICAL CENT

PI Korenberg JK;

WP1: 1998-271791/24.

New isolated Down's syndrome-cell adhesion molecule - used to

developmental and neurological abnormalities

Example 2; page 79; 109pp; English

XX This polynucleotide comprises cDNA fragment E51 that was isolated from a human trisomy 21 foetal brain (14-wk gestation). cDNA CC from a human trisomy 21 foetal brain (14-wk gestation). cDNA CC library following a modified direct cDNA selection technique. CC applied to bacterial and pl artificial chromosomes between EHS2 CC and MX1. E51 was used as a probe to screen the trisomy 21 foetal CC brain library. 62 clones were isolated from the 2 million clones CC of the original library. Overlapping clones were obtained that CC were used to deduce a full-length sequence (see AMV1981) coding

CC for novel Down syndrome-cell adhesion molecule DS-CAM1 (see  
CC AAM42086). A splice variant, DS-CAM2 (see AAV31'88), was also  
CC identified. DS-CAM polypeptides are associated with developmental  
CC and neurological processes. The polypeptides and nucleic acids  
CC are used to develop products for the detection, diagnosis and  
XX therapy of developmental and neurological abnormalities.

50 Sequence 384 BP; 83 A; 119 C; 95 G; 91 T; 0 other;

Query Match	100.0%	Score 20;	DB 19;	Length 388;
Best Local Similarity	100.0%	Ident. No. 5.3;		
Matches	20;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0

```
QY      1 cctgtatgaccctgcaggaac 20
          |||
Db     331 CCTGTATGACCTGCAGGAAC 312
```

RESULT	3
AAV31981	
ID	AAV31981 standard; cDNA: 6604 BP

AC AAV31981.2

DT. 28-SEP-1998 (first entry)

DE · Human Down syndrome-cell adhesion molecule DS-CAM1 CDNA

DS-CAM1; Down syndrome cell adhesion molecule; mental retardation; neuronal transduction; Trisomy 21; mental retardation;

schizencephaly; diagnosis; assay; human; ds; ss  
 holoprosencephaly; colpitis (callosum agenesia);  
 KM

XX  
XX  
XX Homo sapiens.

Key	Location/Qualifiers
XX	
FH	

FT 1000

**FIT**

/\*lag=0

PN W09817795-A1

FD 30-APK-1998.

PF 23-(C)(T)-1997; 97W01US19971.

PR 25-OCT-1996; 9605-0029322.

PA (CEDA-) CEDARS SINAI MEDICAL CENT

PI Korenberg JK  
xx

DR WPT; 1998-271791/24  
DP P-PS10B: AAW42086.

XX  
XX  
PT  
New isolated power

PT developmental and neurological abnormalities

XX  
PS Claim 3; page 65-72; 109pp; English

XX This cDNA clone codes for down syndrome-cell adhesion molecule  
CC DS-CAM1 (see AAW42086), a cell surface glycoprotein belonging to a  
CC novel subclass of the Ig superfamily with highest homology to  
CC neutral cell adhesion molecules. A modified direct cDNA selection  
CC technique was applied to bacterial and pl artificial chromosomes  
CC between EHS2 and MX1 by using cDNA from trisomy 21 human foetal  
CC brain. A unique cDNA fragment, designated E51 (see AAW31982), was  
CC obtained and used to screen a trisomy 21 human foetal brain (14-wk

gestation) cDNA library. Further clones were obtained by exon trapping. The complete DS-CAM1 cDNA sequence was deduced from overlapping clones. A splice variant cDNA (see AAV31988), encoding non-membrane bound DS-CAM2 (see AAM42087), was also identified. The DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome 21q22.2-22.3. The invention also provides murine DS-CAM partial sequences (see AAV31985-87), expression vectors and host cells, transgenic animals, antisense oligonucleotides, and primers useful for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are associated with developmental and neurological processes. They can be used in e.g. neural prosthetic devices used in entubulation methods of repairing (regenerating) damaged or severed peripheral nerves. The products can also be used in detection, diagnosis and therapy of developmental and neurological abnormalities such as Down syndrome, mental retardation, holoprosencephaly, agenesis of the corpus callosum, or schizencephaly. Antisense oligonucleotides are used for inhibiting translation of mRNA.

Sequence 6604 BP; 1673 A; 1827 G; 1336 T; 0 other:

Query Match 100.0%; Score 20; DB 19; Length 6604;  
Best Local Similarity 100.0%; Pred. NO. 6.2; Mismatches 0; Gaps 0;  
Matches 20; Conservative 0; Indels 0; Caps 0;

1 cctgatagcctgcaggaag 20  
|||||  
5054 cctgatagcctgcaggaag 5073

RESULT 4

AAV31988 standard; cDNA: 6413 BP.

AAV31988:

28-SEP-1998 (first entry)

Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.

DS-CAM2: Down syndrome-cell adhesion molecule; neural cell;  
signal transduction; trisomy 21; mental retardation;  
holoprosencephaly; corpus callosum agenesis;  
schizencephaly; diagnosis; assay; human; ds; ss.

Homo sapiens.

Key location/Qualifiers  
CDS 453..5168  
/\*tag- a

MO9817795-A1.

30-APR-1998.

23-OCT-1997; 97WO-0519547.

25-OCT-1996; 96US-0029322.

(CEDARS SINAI MEDICAL CENT.

Korenberg JR;

WPI: 1998-271791/24.

P-PSDH; AAM42087.

New isolated Down's Syndrome-cell adhesion molecule - used to develop products for detection, diagnosis and therapy of developmental and neurological abnormalities

Claim 3: Page 83-90; 109pp; English.

This cDNA clone codes for Down syndrome-cell adhesion molecule

DS-CAM2 (see AAM42087), an extracellular, soluble protein belonging to a novel subclass of the Ig superfamily with highest homology to neural cell adhesion molecules. A modified direct cDNA selection technique was applied to bacterial and plasmid artificial chromosomes between EFS2 and XW1 by using cDNA from trisomy 21 human foetal brain. A unique cDNA fragment, designated E51 (see AAV31982), was obtained and used to screen a trisomy 21 human foetal brain (14-wk gestation) cDNA library. Further clones were obtained by exon trapping, and the complete DS-CAM2 cDNA sequence was identified. DS-CAM2 is a splice variant of membrane-bound DS-CAM1 (see AAM42086) and is deleted for the entire transmembrane domain. The DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome 21q22.2-22.3. The invention also provides murine DS-CAM partial sequences (see AAV31985-87), expression vectors and host cells, transgenic animals, antisense oligonucleotides, and primers useful for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are associated with developmental and neurological processes. They can be used in e.g. neural prosthetic devices used in entubulation methods of repairing (regenerating) damaged or severed peripheral nerves. The products can also be used in detection, diagnosis and therapy of developmental and neurological abnormalities such as Down syndrome, mental retardation, holoprosencephaly, agenesis of the corpus callosum, or schizencephaly. Antisense oligonucleotides are used for inhibiting translation of mRNA.

Sequence 6413 BP; 1633 A; 1781 G; 1707 G; 1292 T; 0 other:

Query Match 92.0%; Score 18.4; DB 19; Length 6413;  
Best Local Similarity 95.0%; Pred. NO. 32;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

1 cctgatagcctgcaggaag 20  
|||||  
5054 cctgatagcctgcaggaag 5073

RESULT 5

AAV31987 standard; cDNA: 2173 BP.

AAV31987:

28-SEP-1998 (first entry)

Mouse Down syndrome-cell adhesion molecule DS-CAM 3' cDNA.

DS-CAM: Down syndrome-cell adhesion molecule; neural cell;  
signal transduction; trisomy 21; mental retardation;  
holoprosencephaly; corpus callosum agenesis;  
schizencephaly; diagnosis; assay; mouse; ds; ss.

Mus sp.

MO9817795-A1.

30-APR-1998.

23-OCT-1997; 97WO-0519547.

25-OCT-1996; 96US-0029322.

(CEDARS SINAI MEDICAL CENT.

Korenberg JR;

WPI: 1998-271791/24.

New isolated Down's Syndrome-cell adhesion molecule - used to develop products for detection, diagnosis and therapy of developmental and neurological abnormalities

Claim 2: Page 81-83; 109pp; English.

XX This cDNA sequence comprises the 3' region of a cDNA clone  
CC for murine Down syndrome-cell adhesion molecule (DS-CAM), a  
CC number of a novel subclass of the Ig superfamily with homology to  
CC neutral cell adhesion molecules. The middle region (see AAV31986)  
CC and 5' region (see AAV31985) of the clone are also provided. The  
CC murine DS-CAM clone was isolated from a C57 Black/6 mouse brain  
CC cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31988) as  
CC probe. The invention also provides human DS-CAM proteins (see  
CC AAV42086-87), as well as expression vectors and host cells,  
CC transgenic animals, antisense oligonucleotides, and primers useful  
CC for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are  
CC associated with developmental and neurological processes. They can  
CC be used in e.g. neural prosthetic devices used in encephalation  
CC methods of repairing (regenerating) damaged or severed peripheral  
CC nerves. The products can also be used in detection, diagnosis and  
CC therapy of developmental and neurological abnormalities such as  
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
CC are used for inhibiting translation of mRNA.

9  
X  
SQ Sequence 2173 BP: 585 A: 598 C: 554 G: 436 T: 0 other:

Query Match 87.0%: Score 17.4: DB 19: Length 2173:  
Best Local Similarity 94.7%: Pred. No. 83: 1: Indels 0: Gaps 0:  
Matches 18: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

UY 2 cctatgacctgcaagaa 20  
|||||  
Db 658 cctgatacctgcaagaa 676

RESULT 6  
ID AAA11722 standard: cDNA: 1438 BP.  
XX  
AC AAA11722:  
XX  
XX 14-JUL-2000 (first entry)  
XX  
DE Human CSCP-2 cDNA.  
XX  
KW human: CSCP-1; cell surface glycoprotein; CSCP-2; anlaemic;  
KW neurotrophic; neuroprotective; anticonvulsant; gene therapy; treatment;  
KW diagnosis; hematologic disorder; anemia; myeloproliferative disorder;  
KW lymphoma; karyotypic disorder; Klinefelter syndrome; Turner syndrome;  
KW neuronal disorder; akathisia; Alzheimer's disease; amnesia; calanomia;  
KW epilepsy; neurofibromatosis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 84..1400  
FT /\*tag= a  
FT /\*product= "CSCP-2"

XX  
XX US6043056-A.  
XX  
XX 28-MAR-2000.  
XX  
XX 06-NOV-1998: 9805-0187331.  
XX  
XX 06-NOV-1998: 9805-0187331.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Yue II, Goryone GA, Baughm MR, Corley NC, Guegler KJ;  
XX WPI: 2000-270341/23.  
XX P-PSDB; AAW90962.  
XX  
XX Polynucleotide encoding cell surface glycoproteins is useful for

PT diagnosis, treatment and prevention of hematologic disorders,  
XX karyotypic disorders and neuronal disorders  
XX  
XX Claim 5: Fig 3A-D: 38pp: English.

CC This invention describes the novel human cell surface glycoproteins  
CC CSCP-1 and CSCP-2. The products of the invention have anlaemic,  
CC neurotrophic, neuroprotective and anticonvulsant activity and can be used  
CC for gene therapy. The encoding nucleic acids are useful for the  
CC diagnosis, treatment and prevention of hematologic disorders such as  
CC anemia, myeloproliferative disorders and lymphoma, karyotypic disorders  
CC such as Klinefelter syndrome and Turner syndrome, and neuronal  
CC disorders such as akathisia, Alzheimer's disease, amnesia, calanomia,  
CC epilepsy and neurofibromatosis. The products of the invention are also  
CC useful as probes for mapping the gene sequences encoding CSCP-1 and  
CC CSCP-2. This sequence encodes the human CSCP-2 protein described in  
CC the method of the invention.

SQ Sequence 1438 BP: 355 A: 369 C: 385 G: 329 T: 0 other:

Query Match 84.0%: Score 16.8: DB 21: Length 1438:  
Best Local Similarity 90.0%: Pred. No. 1,5e+02: 2: Indels 0: Gaps 0:  
Matches 18: conservative 0: Mismatches 2: Indels 0: Gaps 0:

OY 1 cctatgacctgcaagaa 20  
|||||  
Db 863 cctgatacctgcaagaa 882

RESULT 7  
ID AAA96337 standard: cDNA: 1579 BP.  
XX  
XX AAA96337:  
XX  
XX 08-FEB-2001 (first entry)  
XX  
DE cDNA encoding a novel 1 typepeptide designated PRO4334.  
XX  
KW Secreted protein; Transmembrane protein; PRO1484; PRO4334; PRO1122;  
KW PRO1889; PRO1890; PRO1887; PRO4353; PRO4357; PRO4405; PRO4356;  
KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO6030;  
KW PRO4424; PRO4422; PRO4430; PRO4499; tumor; obesity; diabetes;  
KW Insulinoma; kidney disorder; kidney disease; nephropathy;  
KW Schenlein-Blumreich purpura; ocular disease; dermatitis herpetiformis;  
KW Crohn's disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 84..1406  
FT /\*tag= a  
FT sig\_peptide 84..149  
FT /\*tag= b

XX  
XX WO200056889-A2.  
XX  
XX 28-SEP-2000.  
XX  
XX 01-MAR-2000: 2000WO-0505601.  
XX  
XX 23-MAR-1999: 9905-0125774.  
XX 23-MAR-1999: 9905-0125778.  
XX 24-MAR-1999: 9905-0125826.  
XX 31-MAR-1999: 9905-0127035.  
XX 05-APR-1999: 9905-0127706.  
XX 21-APR-1999: 9905-0130359.  
XX 27-APR-1999: 9905-0131270.  
XX 27-APR-1999: 9905-0131272.  
XX 27-APR-1999: 9905-0131291.  
XX 04-MAY-1999: 9905-0132371.  
XX 04-MAY-1999: 9905-0132379.





```

XX
DR WPI: 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta.
XX
PS Claim 25: SEQ ID No 3938; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 1954 BP; 634 A; 350 C; 447 G; 523 T; 0 other:

Query Match          84.0%; Score 16.8; DB 22; Length 1954;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

{
  1 cctgtatgacctgcagaag 20
    |||
  200 ccagatgacctgcagaag 219

Db
}

RESULT 10
AA103753
ID AA103753 standard; DNA; 1954 BP.
XX
XX AA103753;
XX
XX 09-OCT-2001 (first entry)
XX
XX Probe #3744 used to measure gene expression in human breast sample.
DE
XX
XX Probe: human; breast disease; breast cancer; development disorder; ss;
KM Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
OS
XX
XX MO200157270-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 29-JAN-2001; 2001WO-US00661.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
H
XX
XX 26-MAY-2000; 2000US-0207456.
H
XX
XX 30-JUN-2000; 2000US-0608408.
PH
XX
XX 03-AUG-2000; 2000US-0632366.
PR
XX
XX 21-SEP-2000; 2000US-0234687.
PR
XX
XX 27-SEP-2000; 2000US-0236359.
PR
XX
XX 04-OCT-2000; 2000GB-0024263.
PA
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DK;
PI
XX
XX WPI: 2001-476286/51.
DR
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast.
XX
XX Claim 25: SEQ ID No 3744; 322pp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases

```

```

CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1954 BP; 634 A; 350 C; 447 G; 523 T; 0 other:

Query Match          84.0%; Score 16.8; DB 22; Length 1954;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

{
  1 cctgtatgacctgcagaag 20
    |||
  200 ccagatgacctgcagaag 219

Db
}

RESULT 11
AA114982
ID AA114982 standard; DNA; 1955 BP.
XX
XX AA114982;
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #4915 for gene expression analysis in human cervical cell sample.
DE
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157278-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX
XX 26-MAY-2000; 2000US-0207456.
PR
XX
XX 30-JUN-2000; 2000US-0608408.
PR
XX
XX 03-AUG-2000; 2000US-0642466.
PR
XX
XX 21-SEP-2000; 2000US-0234687.
PR
XX
XX 27-SEP-2000; 2000US-0236359.
PR
XX
XX 04-OCT-2000; 2000GB-0024263.
PA
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DK;
PI
XX
XX WPI: 2001-488901/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells.
XX
XX Claim 25: SEQ ID No 4915; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other:

```

Query Match 84.0%; Score 16.8; DB 22; Length 1955;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cctgtatgacctgcagaag 20  
 || |||||  
 db 1641 ccagtatgacctgcagaag 1660

## RESULT 12

AA136328  
 ID AA136328 standard; DNA: 1955 BP.

XX AA136328;

XX 17-OCT-2001 (first entry)

DE Probe #5014 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
 XX genetic disorder; ss.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SC, Hanzel DK, Chen W, Rank DK;

XX WPI: 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 25; SEQ ID No 5014; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

XX The present sequence is one such probe. The probes are useful for

XX producing a microarray for predicting, measuring and displaying gene

XX expression in samples derived from human placenta. The probes are useful

XX for antenatal diagnosis of human genetic disorders.

XX Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other;

XX Query Match 84.0%; Score 16.8; DB 22; Length 1955;

XX Best Local Similarity 90.0%; Pred. No. 1.5e+02;

XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## RESULT 13

AA104749  
 ID AA104749 standard; DNA: 1955 BP.

XX AA104749;

XX

DT 09-OCT-2001 (first entry)

XX Probe #4740 used to measure gene expression in human breast sample.

DE Probe; human; breast disease; breast cancer; development disorder; ss;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US00661.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SC, Hanzel DK, Chen W, Rank DK;

XX WPI: 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression

XX in a human breast -

XX Claim 25; SEQ ID No 4740; 322bp; English.

XX The present invention relates to novel single exon nucleic acid probes.

XX The present sequence is one such probe. The probes are useful for

XX measuring human gene expression in a human breast sample, where the probe

XX hybridises at high stringency to a nucleic acid expressed in the human

XX breast. The probes are useful for predicting, diagnosing, grading,

XX staging, monitoring and prognosing diseases of the human breast,

XX particularly those diseases with polygenic aetiology. The diseases

XX include: breast cancer, disorders of development, inflammatory diseases

XX of the breast, fibrocystic changes, proliferative breast disease and

XX non-carcinoma tumours.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at [http://wipo.int/patlib/published\\_pat\\_sequences](http://wipo.int/patlib/published_pat_sequences).

XX Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other;

XX Query Match 84.0%; Score 16.8; DB 22; Length 1955;

XX Best Local Similarity 90.0%; Pred. No. 1.5e+02;

XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## RESULT 14

AAH57390  
 ID AAH57390 standard; cDNA: 6016 BP.

XX AAH57390;

XX 10-SEP-2001 (first entry)

XX Human skeletal muscle cell specific cDNA sequence SEQ ID NO:230.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;

XX lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;

XX metabolic disease; developmental disease; cytostatic; immunomodulatory;

XX

KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.  
 XX Homo sapiens.  
 OS  
 XX MO200132927-A2.  
 PN  
 XX 10-MAY-2001.  
 PD  
 XX 02-NOV-2000; 2000MO-US0396.  
 PF  
 XX 04-NOV-1999; 99US-0163508.  
 PR  
 XX (INCYTE GENOMICS INC.  
 PA  
 P1 Sornasse T, Seillhamer JJ, Watson GA;  
 P2 WPI; 2001-291057/30.  
 DR  
 XX New cell and tissue specific polynucleotides useful for diagnosis,  
 PT prognosis or monitoring of treatments for disorders where the gene is  
 P1 associated with a cancer, immunopathology or neuropathology -  
 PS  
 XX Claim 1; Page 158-159; 327pp; English.  
 CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide  
 CC sequences (1). (1) can have cytosolic, immunomodulatory and  
 CC neuroprotective activities, and can be used in gene therapy. (1) and  
 CC proteins (11) encoded by them are used in high throughput screening  
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or  
 CC their fragments, immunoglobulins, inhibitors, drug compounds and  
 CC pharmaceutical agents. Expression of (1) in a sample indicates the  
 CC differentiation of embryonic stem cells into a tissue selected from  
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
 CC tissues. (1) and (11) are used to produce an expression profile that  
 CC defines a metabolic or developmental process, treatment, condition,  
 CC disease or disorder. The gene profile can be used for diagnosis,  
 CC prognosis or monitoring of treatments and for investigating a  
 CC predisposition to a disorder where the gene is associated with a  
 CC cancer, immunopathology or neuropathology.  
 CC  
 XX  
 SQ Sequence 6016 BP; 1953 A; 1297 C; 1590 G; 1176 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 6016;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q 1 cctatagacttgcagaaq 20  
 ||| ||||| ||| ||  
 Db 3100 cctgagatgacctgcagcagcag 3119

RESULT 15  
 AA224616  
 ID AA224616 standard; cDNA; 8948 BP.  
 XX  
 AC AA224616;  
 XX  
 UF 07-DEC-1999 (first entry)  
 XX  
 DE Human lung tumor associated polynucleotide.  
 XX  
 KW Human; lung tumor; lung cancer; T cell stimulation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9947674-A2.  
 PN  
 XX 23-SEP-1999.  
 PD  
 XX 17-MAR-1999; 99WO-US05798.  
 PF  
 XX

PR 18-MAR-1998; 98US-0040802.  
 PR 18-MAR-1998; 98US-0040984.  
 PR 27-JUL-1998; 98US-012912.  
 PR 27-JUL-1998; 98US-012933.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Reed SG, Wang T;  
 PI  
 XX WPI; 1999-571839/48.  
 DR  
 XX New isolated lung tumor polynucleotides, used to develop products for  
 PT the treatment, prevention and monitoring the progression of lung cancer  
 P1  
 PS  
 XX Claim 12; Page 111-114; 148pp; English.  
 CC  
 CC The invention provides isolated human lung tumor nucleic acids and  
 CC polypeptides. The polypeptides can be used for the treatment of lung  
 CC cancer. The polypeptides and polynucleotides can be used to stimulate T  
 CC cells or antigen presenting cells for use in the treatment of lung  
 CC cancer. The polypeptides and monoclonal antibodies specific for the  
 CC polypeptides can also be used to inhibit the development of lung cancer.  
 CC Agents which bind the polypeptides can be used for detecting lung cancer  
 CC and for monitoring the progression of lung cancer.  
 CC  
 XX  
 SQ Sequence 8948 BP; 2786 A; 2037 C; 2367 G; 1758 T; 0 other;

Query Match 84.0%; Score 16.8; DB 20; Length 8948;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q 1 cctatagacttgcagaaq 20  
 ||| ||||| ||| ||  
 Db 4862 ccaatagacttgcagaaq 4881

Search completed: March 4, 2002, 05:21:26  
 Job time: 17515 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:11:33 ; Search time 226.99 Seconds  
(without alignments)  
19,955 Million cell updates/sec

Title: US-08-956-991-6

Sequence: 1 cctglatgacctgcaggaag 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database: Issued\_Patents.NA:\*

1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCRTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	1438	US-09-187-331-4	Sequence 4, Appl 1
2	15.8	79.0	840	US-08-998-416-530	Sequence 530, Appl 1
3	15.8	79.0	1296	US-08-090-523-1	Sequence 1, Appl 1
4	15.8	79.0	1296	US-08-090-523-3	Sequence 3, Appl 1
5	15.8	79.0	1296	US-08-398-627-1	Sequence 1, Appl 1
6	15.8	79.0	1296	US-08-398-627-3	Sequence 3, Appl 1
7	15.8	79.0	1296	US-08-406-858-1	Sequence 1, Appl 1
8	15.8	79.0	1296	US-08-406-858-3	Sequence 3, Appl 1
9	15.8	79.0	1296	PCT-US91-04036-1	Sequence 1, Appl 1
10	15.8	79.0	1296	PCT-US91-04036-3	Sequence 3, Appl 1
11	15.8	79.0	1296	PCT-US94-05275-1	Sequence 1, Appl 1
12	15.8	79.0	1296	PCT-US94-05275-3	Sequence 3, Appl 1
13	15.8	79.0	1323	US-07-735-065-3	Sequence 3, Appl 1
14	15.8	79.0	1323	US-08-469-202-13	Sequence 13, Appl 1
15	15.8	79.0	1323	US-08-464-414C-13	Sequence 13, Appl 1
16	15.8	79.0	1323	US-08-229-515A-14	Sequence 14, Appl 1
17	15.8	79.0	3955	US-08-645-865-14	Sequence 14, Appl 1
18	15.4	77.0	24	PCT-US93-00321-2	Sequence 2, Appl 1
19	15.4	77.0	24	PCT-US93-00321-4	Sequence 4, Appl 1
20	15.4	77.0	5102	US-08-494-168-1	Sequence 1, Appl 1
21	15.2	76.0	22	US-08-184-422-15	Sequence 15, Appl 1
22	15.2	76.0	22	US-08-589-771B-15	Sequence 15, Appl 1
23	15.2	76.0	671	US-09-182-117-26	Sequence 26, Appl 1
24	15.2	76.0	818	US-09-363-970-39	Sequence 39, Appl 1
25	15.2	76.0	1003	US-09-363-970-38	Sequence 38, Appl 1
26	15.2	76.0	1186	US-08-064-121-2	Sequence 2, Appl 1
27	15.2	76.0	1186	US-08-478-015-2	Sequence 2, Appl 1

28	15.2	76.0	1186	US-08-475-975-2	Sequence 2, Appl 1
29	15.2	76.0	1186	US-09-084-889-2	Sequence 2, Appl 1
30	15.2	76.0	1415	US-08-257-341-8	Sequence 8, Appl 1
31	15.2	76.0	1515	US-09-080-625-6	Sequence 6, Appl 1
32	15.2	76.0	1521	US-08-646-538-25	Sequence 25, Appl 1
33	15.2	76.0	1521	US-08-646-538-31	Sequence 31, Appl 1
34	15.2	76.0	1521	US-09-503-222-25	Sequence 25, Appl 1
35	15.2	76.0	1521	US-09-503-222-31	Sequence 31, Appl 1
36	15.2	76.0	1701	US-09-080-625-1	Sequence 1, Appl 1
37	15.2	76.0	1853	US-08-404-732A-6	Sequence 6, Appl 1
38	15.2	76.0	1872	US-08-422-108-2	Sequence 2, Appl 1
39	15.2	76.0	2119	US-09-032-372-6	Sequence 6, Appl 1
40	15.2	76.0	2120	US-08-404-732A-4	Sequence 4, Appl 1
41	15.2	76.0	2150	US-09-263-023-1	Sequence 1, Appl 1
42	15.2	76.0	2220	US-08-864-224-1	Sequence 1, Appl 1
43	15.2	76.0	2224	US-08-404-732A-8	Sequence 8, Appl 1
44	15.2	76.0	2245	US-09-442-055-3	Sequence 3, Appl 1
45	15.2	76.0	2385	US-09-146-283-3	Sequence 3, Appl 1

#### ALIGNMENTS

```
RESULT 1
US-09-187-331-4
: Sequence 4, Application US/09187331
: Patent No. 6043056
: GENERAL INFORMATION:
: APPLICANT: Yue, Henry
: APPLICANT: Corley, Neil C.
: APPLICANT: Guegler, Karl J.
: APPLICANT: Gorgone, Gina A.
: APPLICANT: Baughn, Mariah R.
: TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
: FILE REFERENCE: PI-0631 US
: CURRENT APPLICATION NUMBER: US/09/187,331
: CURRENT FILING DATE: 1998-11-06
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PERL Program
: SEQ ID NO 4
: LENGTH: 1438
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE: _
: OTHER INFORMATION: 2705267
US-09-187-331-4

Query Match      84.0%  Score 16.8:  DB 3:  Length 1438:
Best Local Similarity  90.0%:  pred. No. 17:
Matches 18:  Conservative  0:  Mismatches  2:  Indels  0:  Gaps  0:

QY      1 cctglatgacctgcaggaag 20
      ||||| ||||| ||||| |||
Db      863 cctglatgacctgcaggaag 882

RESULT 2
US-08-998-416-530/c
: Sequence 530, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippson, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Reipschling, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSYPPII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
```

ADDRESSEE: No. 6239264rtis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/COC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8589  
INFORMATION FOR SEQ ID NO: 530:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 840 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAC1370UP  
US-08-998-416-530

Query Match 79.0%; Score 15.8; DB 4; Length 840;  
Best Local Similarity 89.5%; Pred. No. 47;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgataagcctgcaggaag 20  
|||||  
DB 607 CTGTATGACCTGCTGCAGG 589

RESULT 3  
US-090-523-1  
Invention 1, Application US/08090523  
Patent No. 5498830  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced starch biosynthesis  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF  
STREET: 700 Chesterfield Parkway No. 5498830th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/090,523  
FILING DATE: 19930712  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-090-523-1

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgataagcctgcaggaag 20  
|||||  
DB 667 CTGTATGACCTGCTGCAGG 685

RESULT 4  
US-08-090-523-1  
Invention 3, Application US/08090523  
Patent No. 5498830  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced starch biosynthesis  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF  
STREET: 700 Chesterfield Parkway No. 5498830th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/090,523  
FILING DATE: 19930712  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-090-523-3

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgtatgacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACTGCTGGAG 685

RESULT 5  
US-08-398-627-1  
Sequence 1, Application US/08398627  
Patent No. 5608149  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced Starch Biosyntheses  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. B44F  
STREET: 700 Chesterfield Parkway No. 5608149th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/398.627  
FILING DATE: 03-MAR-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/090.523  
FILING DATE: 12-JUL-1993  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293

US-08-398-627-1

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgtatgacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACTGCTGGAG 685

RESULT 6  
US-08-398-627-3  
Sequence 3, Application US/08398627  
Patent No. 5608149  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced Starch Biosyntheses  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. B44F  
STREET: 700 Chesterfield Parkway No. 5608149th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/398.627  
FILING DATE: 03-MAR-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/090.523  
FILING DATE: 12-JUL-1993  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-398-627-3

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgtatgacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACTGCTGGAG 685

RESULT 7  
US-08-406-858-1  
Sequence 1, Application US/08406858  
Patent No. 5648249  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
APPLICANT: Zaleski, James C.  
TITLE OF INVENTION: Method of improving the quality of  
TITLE OF INVENTION: Stored Potatoes  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Company, B84F  
STREET: 700 Chesterfield Parkway No. 5648249th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,858  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE: 18-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,155  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10654)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-7286  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
US-08-406-858-1

Query Match 79.0%; Score 15.8; DB 1; Length 196;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctglatgacctgcgaag 20  
||||||| ||||| |||||  
DB 667 CTGTATGACCTGCTGGAAG 685

RESULT 8  
US-08-406-858-3  
Sequence 3, Application US/08406858  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.

APPLICANT: Zaleski, James C.  
TITLE OF INVENTION: Method of improving the quality of  
TITLE OF INVENTION: Stored Potatoes  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Company, B84F  
STREET: 700 Chesterfield Parkway No. 5648249th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,858  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE: 18-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,155  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10654)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-7286  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
US-08-406-858-4

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctglatgacctgcgaag 20  
||||||| ||||| |||||  
DB 667 CTGTATGACCTGCTGGAAG 685

RESULT 9  
PCT-US91-004036-1  
Sequence 1, Application PCT/US9104036  
GENERAL INFORMATION:  
APPLICANT: Kishore, Ganesh M.  
TITLE OF INVENTION: Increased starch content in plants  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Monsanto Co.  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: PCT/US91/04036  
APPLICATION NUMBER: PCT/US91/04036  
FILING DATE: 19910607  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McBride, Thomas P.  
REGISTRATION NUMBER: 32706  
REFERENCE/DOCKET NUMBER: 38-21(10530)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7357  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
PCT-US91-04036-1

Query Match 79.0%; Score 15.8; Db 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
|||||1111111111111111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 10  
PCT-US91-04036-3  
Sequence 3, Application PC/TUS9104036  
GENERAL INFORMATION:  
APPLICANT: Kishore, Ganesh M.  
TITLE OF INVENTION: Increased Starch Content in Plants  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Monsanto Co.  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/04036  
FILING DATE: 19910607  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McBride, Thomas P.  
REGISTRATION NUMBER: 32706  
REFERENCE/DOCKET NUMBER: 38-21(10530)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7357  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..1293  
PCT-US91-04036-3

Query Match 79.0%; Score 15.8; Db 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
|||||1111111111111111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 11  
PCT-US94-05275-1  
Sequence 1, Application PC/US9405275  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Method of Improving the Quality of Stored  
NUMBER OF SEQUENCES: 26  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

Query Match 79.0%; Score 15.8; Db 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
|||||1111111111111111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 12  
PCT-US94-05275-3  
Sequence 3, Application PC/TUS9405275  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Method of Improving the Quality of Stored  
NUMBER OF SEQUENCES: 26  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 08/070155  
FILING DATE: 28-MAY-1993  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
PCT-US94-05275-3

Query Match 79.0%; Score 15.8; DB 5; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

{ 2 ctglatgacctgcaggaag 20  
||||||| ||||| |||||  
db . 667 CTGTATGACTGCTGGAAG 685

RESULT 13  
US-07-735-065-3  
Sequence 3, Application US/07735065  
PATENT No. 5349123  
GENERAL INFORMATION:  
APPLICANT: Christine K. Shewmaker  
APPLICANT: David M. Stalker  
TITLE OF INVENTION: Glycogen Biosynthetic Enzymes  
TITLE OF INVENTION: In Plants  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene, Inc.  
STREET: 1920 Fifth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.7  
SOFTWARE: Microsoft Word 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/735,065  
FILING DATE: 19910724  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/632,383  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: unassigned  
FILING DATE: 16-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Lassen  
REGISTRATION NUMBER: 31,845  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
REFERENCE/DOCKET NUMBER: CGNE 77-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1323 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-07-735-065-3

Query Match 79.0%; Score 15.8; DB 1; Length 1323;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 ctglatgacctgcaggaag 20  
||||||| ||||| |||||  
Db 682 CTGTATGACTGCTGGAAG 700

RESULT 14  
US-08-469-202-13  
Sequence 13, Application US/08469202  
PATENT No. 5750875  
GENERAL INFORMATION:  
APPLICANT: STALKER, DAVID  
TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CALGENE, INC.  
STREET: 1920 FIFTH STREET  
CITY: DAVIS  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,202  
FILING DATE: 6-JUNE-95  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,881  
FILING DATE: 11-FEB-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Lassen  
REGISTRATION NUMBER: 31,845  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
NAME: Carl J. Schvedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 93-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 916-753-6415  
TELEFAX: 916-753-1510  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1323 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
US-08-469-202-13

Query Match 79.0%; Score 15.8; DB 1; Length 1323;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctglatgacctgcaggaag 20  
||||||| ||||| |||||  
Db 682 CTGTATGACTGCTGGAAG 700

RESULT 15  
US-08-484-434C-13  
Sequence 13, Application US/08484434C

; Patent No. 5969214  
; GENERAL INFORMATION:  
; APPLICANT: STALKER, DAVID  
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Calgene LLC  
; STREET: 1920 Fifth Street  
; CITY: Davis  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 95616  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,434C  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/016,881  
; FILING DATE: 11-FEB-1993  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmedler, Carl J.  
; REGISTRATION NUMBER: 36,924  
; REFERENCE/DOCKET NUMBER: 15593/01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 530-792-2265  
; TELEFAX: 530-792-2463  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1323 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME//KEY: CDS  
; LOCATION: 16..1308  
; US-08-484-434C-13

Query Match 79.0%; Score 15.8; DB 2; Length 1323;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

U 2 cgtatgacctcaggaag 20  
|||||  
Db 682 cgtatgacctcaggaag 700

Search completed: March 4, 2002, 05:11:34  
Job time: 16948 sec











```

DE      The nucleotide sequence of the Herpes simplex virus ICP4.
XX
XX      Herpes simplex virus ICP4; HSV ICP4; modulator; apoptosis;
XX      stimulation; inhibition; HSV infection; ss.
XX
XX      Herpes simplex virus.
XX
XX      Key      Location/Qualifiers
XX      CDS      361..4257
XX               /*tag=a
XX               /product= "HSV ICP4 protein"
XX
XX      MO9846637-A2.
XX
XX      22-OCT-1998.
XX
XX      16-APR-1998; 98MO-US07573.
XX
XX      16-APR-1997; 97US-0843659.
XX
XX      (ARCH-) ARCH DEV CORP.
XX
XX      Leopardi R, Roizman B;
XX
XX      WPI; 1998-594519/50.
XX      P-PSDB; AAM80810.
XX
XX      Use of herpes simplex virus U(S)3 polypeptide - for developing
XX      products for modulating apoptosis in cells and for identifying
XX      compounds which act as stimulators or inhibitors of apoptosis
XX
XX      Example 2; Pages 60-63; 85pp; English.
XX
XX      This is the nucleotide sequence of Herpes simplex virus ICP4 used in
XX      the method of the invention as modulators of apoptosis. The method;
XX      and products can be used to identify compounds which modulate
XX      (stimulate or inhibit) apoptosis in cells. They can be used to
XX      immortalise cells for the study of these cells or for growing cells
XX      in large numbers for the productions of proteins. They can also be
XX      used for stimulating apoptosis in cells, e.g. for treating a subject
XX      with a HSV infection.
XX
XX      Sequence 4257 BP; 412 A; 1768 C; 1663 G; 414 T; 0 other;
XX
XX      Query Match      4.7%; Score 39.2; DB 19; Length 4257;
XX      st Local Similarity 62.0%; Pred. No. 0.37;
XX      ches 62; Conservative 0; Mismatches 38; Indels 0; Gaps
XX      {
XX      2 cggagccggagcgcgagcgagcgcaacgcgycggatgaagcgcgcgctgctcgt 61
XX      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX      2229 CAGGCGGGGCGCTCGCGCTCGCGCGCGCGCGCGCGCGGAGACGCGGCTCC 2170
XX      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX      62 cgcgtgcgcgtcgctcgcgcgagcgcgagcgcgagcgagcgagcgagcgag 101
XX      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX      2169 CCGCGCGGGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2130
XX
XX      RESULT 7
XX      AAV10362/c
XX      ID AAV10362 standard; cDNA; 4257 BP.
XX
XX      AAV10362;
XX
XX      30-JUN-1998 (first entry)
XX
XX      Infected cell protein number 4 alpha-4 gene.
XX
XX      Infected cell protein number 4; ICP4; alpha-4; cell apoptosis;
XX      therapeutics; ss.
XX
XX      Herpes simplex virus.
XX

```

XX	Key		Location/Qualifiers
FT	CDS	361..4257 /*tag= a /product= "Infected cell protein"	
FT			
xx			
PN	M09804.709-A2.		
xx			
PD	05-FEB-1998.		
xx			
PE	23-JUL-1997;	97WU-US12904.	
xx			
PR	26-JUL-1996;	96US-U690473.	
xx			
PA	(ARCH-) ARCH DEV CORP.		
PI	Leopardi R,	Koizman B;	
DR	WI:	1998-130697/12.	
DR	P-PsDB:	AAM40200.	
PT			
PS			
CC			
CC	The herpes simplex virus (HSV) alpha-4 gene encodes infected cell protein number 4 (ICP4). Both the alpha-4 gene and ICP4 can be used to block cell apoptosis. Similarly the administration of an agent that inhibits ICP4 or the alpha-4 gene can induce apoptosis in HSV infected cells. This can be used for the immortalisation of cells, production of proteins, gene therapy, or inhibition of cell death induced in vivo. They can also be used for production of therapeutics comprising inhibitors of HSV ICP4 function, useful for treating HSV function.		
SQ	Sequence	4257 BP; 412 A; 1768 C; 1663 G; 414 T; 0 other;	
OY	Query Match	4.7%; Score 39.2; DB 19; Length 4257; Best Local Similarity 62.0%; Pred. No. 0.37; Mismatches 38; Indels 0; Caps 0	
DG	Matches 62; Conservative 0; Mismatches 38; Indels 0; Caps 0		
OY	2 cggcgcgacgccgaagccatcgcaaccgaggcggcggcgcttgactcgtt 61                                     		
DG	2229 CAGCGCAGGCCTGTGGTCGCCGACCGCGCAAGCGGCGCGTCTCGCGA 2170  OY    62 cgcttccttgtttctctcgaagcccccatcacgcag 101                                     DGB 2169 CGTGCGGGAATTTAAATTATTAATACTCATCACCGCG 2130		
RESULT H			
AAZ29335/C	ID	AAZ29335 standard; DNA: 7686 BP.	
AAZ29335;			
DT	25-FEB-2000	(first entry)	
DE	Human acrosome reaction protein-IPKREJ encoding DNA.		
XN	Acrosome reaction protein; p3; sperm protein; polycystin-1; suresl;		
KM	Polycystic kidney disease and receptor for egg jelly protein; PKDREJ;		
KM	Zona pellucida; fertility; contraceptive; gene therapy; ds.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..6762 /*tag= a /product= "Acrosome reaction protein (PKDREJ)"	
FT	polyA_signal	7636..7641 /*tag= b	
FT			









XX	Homo sapiens.
XX	location/Qualifiers
FH	Key
FT	CDS
FT	176..318 /*tag= a
FT	/note= "encodes KA-2 receptor"
FT	sig_peptide
FT	176..229
FT	mat_peptide
FT	230..315 /*tag= c
FN	EP529995-A.
XX	
HD	03-MAR-1993.
XX	
PP	25-AUG-1992:
XX	92EP-0307724.
PR	27-AUG-1991:
XX	9JUS-0750081.
}	(ALLE-) ALLEXIX BIOPHARMACEUTICALS INC.
Pt	Kamboj R, Nutt SL, Shekter L, Wosnick MA;
DR	WPI: 1993-069002/09.
DR	P-PSDB: AAK32259.
XX	
PT	Isolated polynucleotide encoding human excitatory amino acid-2 receptor - useful for determining binding affinity of cpds. for the receptor in assaying for drugs to treat CNS disorders
PS	Claim 1; Fig 1: 29pp: English.
XX	
CC	The sequence is that of the gene encoding the human KA-2 receptor, a kainate high affinity receptor of the EAA2 (Excitatory amino Acid) family. The sequence can be used in the identification of sequence related genes. It can also be used to produce cells able to produce the receptor, these cells and membrane fragments of them can be used to assay compounds for binding affinity to human EAA receptor i.e. to screen compounds for potential use as drugs against CNS disorders.
XX	
SQ	Sequence 3694 BP; 678 A; 1242 C; 1115 G; 658 T; 1 other:
	Query Match            4.3%; Score 36.6; DB 14; Length 3694; Best local Similarity 53.1%; Pred. No. 2; Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
G	2 cgagcgaggcgtgctgcggacgaacgagccgagcgagcgagcgagcgagcgatgtgatactgtctt 61
b	3107 CCGCCACCTCCTCCCGGGGCGCGCGGCAGCGCCGCGGCGGCGGTGTGCTTTCGG 3048
OY	62 cgctgctgcctgctgcgcggagcgagccggaccgagcgagcgagtgtgatactgtctt 121
Dd	3047 CGGGGAGCGCCAGCGCACCGGAGCGCCGCCGCCGCCGAGCGCCGACGCGCTTGATGC 2988
Oy	122 ctcttgtttcagaageltcgaaatgt 148
Dd	2987 GCCGCACTCCTGGCAGACGCGCAGCT 2961
RESULT 15	
AAT63571	
ID	AAT63571 standard; DNA: 1524 BP.
XX	AAT63571;
XX	30-JUN-1997 (first entry)
DT	Xylanase gene obtained by soil DNA amplification.
DE	Xylanase; soil DNA; ds.
XX	
RX	NoL identified.

XX MO9712991-AI.  
PN 10-APR-1997.  
PD  
XX 20-SEP-1996; 96WO-CAM00627.  
PP  
XX 22-SEP-1995; 95US-0004157.  
PR  
XX (TERR-) TERRAGEN DIVERSITY INC.  
PA  
XX Radomski CCA, Secow KT, Warren RAJ, Yap WH;  
PI  
DR WPI: 1997-226234/20.  
DR P-PSDB; AAM09777.  
XX  
PT xylanase gene sequences - obid., by recovering DNA from soil samples  
PP and PCR amplification using primers based on xylanase genes  
XX  
PS Claim 12: Page 24-26: 34pp; English.

XX A DNA sequence (AAT63571) codes for a novel xylanase (AAM09777). It  
CC was isolated from a phenoloid in a soil DNA library using as probe  
CC a gene fragment (AAT63568) obid., by amplification of soil DNA using  
CC degenerate primers (AAT63548-49) based on conserved portions of  
CC family F cellulases. It was sequenced by primer walking over the  
CC expansion insert using the same degenerate primers as initial  
CC extension primers. Subsequent extension primers were constructed  
CC by examining the previously generated sequence data. The novel  
CC xylanase gene can be used for prodn. of xylanase for industrial  
CC applin.

SQ Sequence 1524 BP; 281 A; 537 C; 507 G; 199 T; 0 other;

[illegible]

Search completed: March 4, 2002, 05:22:27  
Job time: 17576 sec











```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-257G14"
/clone.lib="RPCI-24"
/sex="Male"
/cell-type="Spleen/Brain"
/note="Vector: pTRARAC1; Site.1: BamHI; Site.2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTRARAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      222 a      141 c      154 g      193 t
ORIGIN

```

Query Match	17.0%:	Score 142.8:	DB 13:	Length 710:
Best Local Similarity	92.6%:	Pred. No. 1.8e-25:		
Matches 150:	Conservative	0:	Mismatches 12:	Indels 0: Gaps 0
Db	598	ttcaactgtctcagaagctctagatttctcatcacatccacagagagccttgatataaag	657	
Db	521	tttatatctgtttcaggtatctagattttcttatcatcatccacgggacacctgttatattaaag	462	
Oy	658	atgtccaaacaaagatgagctctglacaaactaccgcctgcaatcgccgagacagattctgag	717	
Db	461	atgttcacaaacgaagatggcctgtacaaactaccgcctgtcattccacgggacacagattacacag	402	
Oy	718	ggagagaagagacagagacgaactctgctgagagactctgtctgcaag	759	
Db	401	ggagacacgacgacaaagacacagcgcgacgactgttctgtctacg	360	
RESULT 6				
HSMC18B12				
LOCUS	H.sapiens	DNA	147 bp	GSS
DEFINITION	H.sapiens DNA for trapped exon (ID HMC18B12), genomic survey sequence.			
ACCESSION	X88325			
VERSION	X88325.1		GI:1437728	
KEYWORDS	GSS:			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 147)			
AUTHORS	Chen,H.M., Rossier,C., Chrasl,R. and Antonarakis,S.E.			
JOURNAL	Cloning of trapped exons from human chromosome 21			
TITLE	Unpublished			
RENCE	2 (bases 1 to 147)			
AUTHORS	Antonarakis,S.E.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-MAR-1995) Stylianou E. Antonarakis, Division of Medical Genetics, University and Cantonal Hospital of Geneva, CMU 1 rue Michel-Servet, 1211 Geneva, SWITZERLAND			
REFERENCE	3 (bases 1 to 147)			
AUTHORS	Chen,H., Chrasl,R., Rossier,C., Morris,M.A., Lalliot,M.D. and Antonarakis,S.E.			
TITLE	Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping			
JOURNAL	Genome Res. 6 (8), 747-760 (1996)			
LINE	97011340			
FEATURES	location/Qualifiers			
SOURCE	1..147			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="21"			
	1..147			
	/note="trapped exon"			
BASE COUNT	33 a	33 c	39 g	38 t
ORIGIN				4 others

[illegible]

	RESULT	7	CNSO3BK1/C	GSS	15-MAY-2000
	LOCUS	CNSO3BK1	535 bp DNA		
	DEFINITION	Tetradon nigriviridis genome survey sequence pUC-ori end of clone 01C20 of library G from Tetradon nigriviridis, genomic survey sequence.			
	ACCESSION	AL236602	GI:7895737		
	VERSION	AL236602.1			
	KEYWORDS	GSS; genome survey sequence.			
	SOURCE	Tetradon nigriviridis.			
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodonlinidae; Tetradon.			
	REFERENCE	1 (bases 1 to 535) Roest-Crolius,H., Jalllon,O., Dasilva,C., Fizesmes,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissbach,J. Characterization and repeat analysis of the compact genome of the freshwater putterfish Tetradon nigriviridis			
JOURNAL	TITLE	Unpublished 2 (bases 1 to 535) Roest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizesmes,C., Winker,P., Brotlier,P., Queller,F., Saurin,W. and Weissbach,J. Human gene number estimate provided by genome wide analysis using Tetradon nigriviridis DNA sequence			
REFERENCE	TITLE	Unpublished 3 (bases 1 to 535) Genoscope. Submitted (12-Apr-2000) to the EMBL/genbank/DDBJ databases This sequence is a single-read end was generated as part of a large scale clone-end sequencing project of the Tetradon nigriviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetradon.location/Qualifiers			
COMMENT	FEATURES	1..535 /organism="Tetradon nigriviridis" /db_xref="taxon:99883" /citione="01C20" /citione_1fb="G" /note="Genoscope sequence ID : COBG012BBI0SP1-end ; pUC-ori"			
	BASE COUNT	140 a     107 c     165 g     102 t               21 others			
	ORIGIN				
QY	Query Match	15.0%;	Score 126.6;	DB 13;	Length 535;
	Best Local Similarity	73.0%;	Pred. No. 1.9e-21;		
MATCHES	222:	Conservative	5;	Mismatches	70;
				Indels	7;
				Gaps	5;
QY	148	ttttaaggaagaaccccatcgc--agacctacttgatcaatgatcgctgaagaagt	205		
DN	430	TTTTCTGTGAAGATCTTCACATTCCCAAGCGTTATTTTCATCATGCATCGTGCCAAGAGT	371		





Matches	104	Conservative	1	Mismatches	4	Indels	0	Gaps	0
OY	465	gltttacggagacctatacaqtccglttgtagagctcaaaacatqayagyaatgltc		524					
Db	785	gTTTTCAAGGAGGCCCTACACGCGTGGGGGCGCGATCGCCCTGAGCGGGCAACGTR		726					
OY	525	gcgggtgttcacatgcatatccccctccctcgtgtagagggcagtcgtctgtgtaactatg		584					
Db	725	GCTGTGTTCAAGTGCCATCCCGGCTGCTGTGCACAGGAATATATTCAGCGTGTCTCTCG		666					
OY	585	gagaaagacacgggttcaactgtctcaag		613					
Db	665	AAGAGACACACCTATCTCATCTGTCACAGG		637					

RESULT	12				
CNS02C1Y					
Locus	611 bp	DNA	GSS	12-MAY-2000	
DEFINITION	Tetradon nigroviridis genome survey sequence. pUC-ori end of clone 255N21 of library G from Tetradon nigroviridis, genomic survey sequence.				

ACCESSION	AL191311
VERSION	AL191311.1
KEYWORDS	GSS, genome survey sequence,
SOURCE	Tetradon nigroviridis.
	Tetradon nigroviridis.

ORGANISM  
Eukaryot; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Aclinopterygii; Neopterygii; Teleostei; Euteleostei; Neocleosteii;  
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
Tetraodon; Tetraodontidae; Tetraodon.

REFERENCE  
1 (bases 1 to 611)  
Roest-Crolius, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,  
Bonneau, L., Billault, A., Quelier, F., Saurin, W., Bernot, A. and  
Weissenbach, J.  
1996. Analysis and repeat analysis of the compact genome of the

TITLE	REFERENCE
Characterization and Replication of a freshwater putrefactive Tetraodon nigroviridis Unpublished	2 (basses 1 to 611)
Reest-Crolius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C., Bouvier, F.	

**TITLE** Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL, unpublished  
REFERENCE 3 (bases 1 to 611)  
AUTHORS Genoscope.  
TITLE Direct Submission  
to the EMBL/GenBank/DBJ databases  
(12-PRR-2000)

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetrahodon nitroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetrahodon>.

```

FEATURES
  source      location/Qualifiers
1. .611
   /organism="retrovirus nigroviridis"
   /db_xref="taxon:99883"
   /clone="255N21"

```

BASE COUNT	109 a	161 c	152 q	167 t	22 others
/note="Genoscope sequence ID : C0AG255CG11SP1-end : /note="Genoscope sequence ID : C0AG255CG11SP1-end : pUC-ori"					

ORIGIN	Score 69;	DB 13;	length 611;
Query Match	8.28;		
Circularity	67.58;		
Pred. No.	5.7e-07;		

[illegible]

D<sub>b</sub>    \*462 GTTCTCAGCGAGGACCTACTATTCCAGCCGTGTTTGATCATCACTGGTGGTAACTG  
QY    207 gttgttcgaagacatcgggagaccttgqlygccttgcctccctcgcaugactccctctcttq    266

Db 521 GATTATTCACAGCMCCCGGGGCGCTCGGTGCTCTGTCGCCCGGGGGAAGCGGCCCGCCG 580

Oy 267 a..tctcagatggtacctagcaacgagcgag 297

Db 581 TCGCTCTCGCTGCTGCTCTCGGCCACCGCGGCG 611

RESULT	13
CNS04L51	
LOCUS	656 bp DNA GSS 21-MAY-2000
DEFINITION	CNS04L51 Tetradon nigroviridis genome survey sequence T7 end of clone
	Library of Tetradon nigroviridis, genomic survey

ACCESSION	AJ295695
VERSION	AJ295695.1
KEYWORDS	GI:8044275
SOURCE	GST; genome survey sequence.
ORGANISM	Tetradon nigroviridis Tetradon nigroviridis Vertebrata: Craniata: Vertebrata: Euteleostomi;

REFERENCE  
Bost-Strouls H., de Jallou, O., basilla, C., Piamas, C., Fisher, C.,  
1 (bases 1 to 656)

**AUTHORS** Korsch, C., Hillmann, A., Queller, F., Saurin, W., Bernot, A. and Homedahl, L.  
**TITLE** Characterization and repeat analysis of the compact genome of the freshwater pufferfish *Tetraodon nigroviridis*  
**JOURNAL** Unpublished

REFERENCE	AUTHORS	TITLE
2 (bases 1 to 655)	Rocsi-Crollius, H., Jallion, O., Iasilva, C., Bouneau, L., Fisher, C., Bernol, A., Fizames, C., Winkler, P., Brotier, P., Queller, F., Saurin, W. and Weissbach, J.	Human gene number estimate provided by genome wide analysis using cDNA sequence

**JOURNAL** Unpublished  
**REFERENCE** 3 (bases 1 to 656)  
**AUTHORS** Genoscope.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (12-Apr 2000) to the EMBL/Genbank/DBJ databases of a large

**COMMENT**  
This sequence is a single read and was generated as part of a large scale clone and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES	SOURCE
Location/Qualifiers	
1. .656	
/organism "Nectomon nigroviridis"	
/dt xref "taxon:39884"	
/clone "117118"	

BASE COUNT	211 d	140 c	141 g	164 t	CONC(17)PC91.P1 end : T7
ORIGIN					

Query Match	7.88	Score 66	DB 13	Length 656
Best Local Similarity	69.28	Pred. No. 4,3e-06		
Best Global Similarity	0	Mismatches 40	Indels 0	Caps 0

	Matches	90, conserved	
OY	625	tcaatcattccatcgaggaacttattalataaaqatqltlcaagaacgaatggygcgtlaca	684
Dh	4	ttcttgaccattctttttggagtcaccttcttatattatcttcacaaattgcagaaagaaacgccctctgtcca	63

685 actactcattcaatctccacatggtatcatcttcgcggcgagacagagacaactctgcgca 744  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY  
Db 64 CCTACGCGCTGCATCACAAGACCAACTACACGCGAGAGAGACTGCCCCAGACCACAGGAGCAA 123

QY	745	gaclatlcqt	754
Db	124	GACTCTCTCT	133







[illegible][illegible]





CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 05/08/459,214  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/197792  
FILING DATE: 17-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/958414  
FILING DATE: 08-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/744207  
FILING DATE: 12-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/215466  
FILING DATE: 05-JUL-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/906729  
FILING DATE: 31-DEC-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/827710  
FILING DATE: 07-FEB-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/783910  
FILING DATE: 03-OCT-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 297P206  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3588 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-459-214-32

RESULT 9  
US-08-716-942-24

```

: Sequence 24, application US/08716942
: Patent No. 5849491
: GENERAL INFORMATION:
: APPLICANT: Terragen Diversity Inc.
: APPLICANT: Radomski, Christopher C. A.
: APPLICANT: Seow, Kah Tong
: APPLICANT: Warren, R. Anthony J.
: APPLICANT: Yap, Wei Ho
: TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE
: TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD
: TITLE OF INVENTION: COMPOSITIONS OBTAINED THEREBY
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Updeahl & Larson
: STREET: 1992 Commerce Street, Suite 309
: CITY: Yorktown Heights
: STATE: NY
: COUNTRY: USA
: ZIP: 10598-4412
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS 5.0
: SOFTWARE: Word Perfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/716,942
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/004,157
: FILING DATE: 20-Sep-95
: ATTORNEY/AGENT INFORMATION:
: NAME: Martina T. Larson
: REGISTRATION NUMBER: 32,038
: REFERENCE/PACKET NUMBER: TERR-P-001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (914) 245-9252
: TELEFAX: (914) 962-4330
: TELEX:
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1524
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: HYPOTHEICAL: no
: ANTI-SENSE: no
: FRAGMENT TYPE: internal
: ORIGINAL SOURCE:
: ORGANISM:
: FEATURE:
: NAME/KEY: sequence of xylanase gene identified by
: NAME/KEY: amplification of xylanase fragments from soil
: US-08-716-942-24

```

[illegible]







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:50 ; Search time 6206.58 Seconds  
(without alignments)  
1554.755 Million cell updates/sec

Title: US-08-956-991-8  
Perfect score: 898  
Sequence: 1 tgcggccggttcgaagcct.....gcgcacacatcgttcgtga 898

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
22703874

( 1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estro:\*  
7: em\_estov:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qd\_est1:\*  
11: qd\_est2:\*  
12: qd\_hic:\*  
13: qd\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Mat h	Length	DB ID	Description
C 1	224.4	25.0	279	10	BB336040 BB336040
C 2	223.8	24.9	291	10	BB333216 BB333216
C 3	218.4	24.3	252	10	BB381728 BB381728
C 4	214.8	23.7	270	10	BB336019 BB336019
C 5	212.8	23.7	296	10	BB335137 BB335137
C 6	212.8	23.7	355	10	BB461534 BB461534
C 7	204.4	22.7	247	10	BB330789 BB330789
C 8	192.4	21.4	257	10	BB329941 BB329941
C 9	190.6	21.2	303	10	BB330426 BB330426
C 10	189.2	21.1	354	10	BB331552 BB331552
C 11	189.2	21.0	286	10	BB280255 BB280255
C 12	189.2	21.0	286	10	BB333910 BB333910

C 13	186	20.7	499	11	B1185857
C 14	184.6	20.6	280	10	BB331669
C 15	180	20.0	252	10	BB334670
C 16	175.8	19.6	239	10	BB332000
C 17	174.4	19.4	272	10	BB330770
C 18	159.8	17.8	216	10	BB328887
C 19	146.2	16.3	219	10	BB342624
C 20	137.8	15.3	212	10	BB330864
C 21	128.4	14.3	194	10	BB331191
C 22	122.4	13.6	287	11	BP928859
C 23	117	13.0	163	10	BB330436
C 24	111.6	12.4	229	10	BB333024
C 25	100.8	11.2	1013	13	CNS02AK7
C 26	92.6	10.3	965	13	CNS02WH1
C 27	64.6	7.2	643	13	AZ790921
C 28	63.8	7.1	794	13	CNS03C7U
C 29	58	6.5	605	13	FR0021994
C 30	58	6.4	938	13	FR0022042
C 31	57.4	6.4	443	13	CNS03410
C 32	55.6	6.2	443	13	FR0022007
C 33	54.4	6.1	385	13	CNS04LEK
C 34	52.6	5.9	499	13	FR0022008
C 35	51	5.7	1101	13	CNS057K8
C 36	50	5.6	564	13	FR0030722
C 37	45	5.0	1101	13	CNS00KR2
C 38	44.4	4.9	73	13	CNS03K9Y
C 39	43.4	4.8	430	13	FR0022072
C 40	43	4.8	419	10	A1750970
C 41	41.8	4.7	405	10	A1005408
C 42	41.8	4.7	969	13	CNS05MTE
C 43	40.8	4.5	573	11	BF266074
C 44	39.6	4.4	930	10	A1566074
C 45	38.8	4.3	423	11	BC263635

#### ALIGNMENTS

RESULT 1  
LOCUS BB336040/c  
DEFINITION BB336040 RIKEN full-length enriched, 10 days neonate medulla oblongata Mus musculus cDNA clone B83046F24 3' similar to AF217525 Homo Sapiens c10nc cDS1 Down syndrome cell adhesion molecule (USCAM) mRNA, mRNA sequence.

ACCESSION BB336040.1 GI:9044803  
VERSION BB336040.1  
KEYWORDS  
SOURCE house mouse  
ORGANISM Mus musculus

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

BB336040 279 bp mRNA EST 11-JUL-2000  
BB336040 RIKEN full-length enriched, 10 days neonate medulla oblongata Mus musculus cDNA clone B83046F24 3' similar to AF217525 Homo Sapiens c10nc cDS1 Down syndrome cell adhesion molecule (USCAM) mRNA, mRNA sequence.  
BB336040  
BB336040.1 GI:9044803  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 279)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Aikawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hata, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, D., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takehashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamane, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222



BASE COUNT	81 a	87 c	59 g	64 l
ORIGIN				

ov 447 cactatcalagcctgtgālqqcctgttgcacatcaaacctcatalggtccaccctgagcgrra 500

507 tggcgcgagccltggacalcclctggalccccaalcccgaggga.....

[illegible][illegible][illegible]

DEFINITION	BB381/28 KIKEN IUII
------------	---------------------

ALLEGATION  
DO 101720  
0107AE  
010720

Chordata: Vertebrata: Euteleostei

Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Ogasawara, S., and Tanaka, K. (1997) The effect of the

Tsunoda, I., Watanabe, Y., Watanabe, Y., Watanabe, Y.

JOURNAL OF  
COMPARATIVE  
POLITICAL  
ECONOMY

Tsurumi-ku, Yokohama, Kanagawa 230-0045

Source	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

100

site 1: BamHI; site 2: BamHI; CDNA library was

RIKEN. Division of Experimental Animal Research, 2-1-39

prepared by using crenatose enzyme

CTTAAATTCACCCCGGCGGGGAGCCTTAATTAAATTAATCAATTTAAT

Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

```

Source: Matlab;
Script: 218-4;
PB: 10;
Length: 252;

```

447 GactgttGatggcctgtgtatgttggcctggttccaccctgagctgga 500

50/ tggctgcgagccttgcacatctcttgatcccaatcccgaggccccc-111  
QY

2019-2020

[illegible]























GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:22:27 : Search time 578.64 Seconds  
(without alignments)  
1330.496 Million cell updates/sec

Title: US-08-956-991-8  
898  
Sequence: 1 tgcgcgcgttgcacgctt.....gcgccacatcgttgcctga 898

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 930021 seqs, 428662619 residues

1 number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq.1101.\*  
1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	898	100.0	898	19	AAV31986
2	502	55.9	6413	19	AAV31988
3	502	55.9	6604	19	AAV31981
4	244.8	27.3	5824	22	AA158196
5	43	4.8	5720	22	AA158196
6	38.8	4.3	1225	18	AA158196
7	38.8	4.3	1228	18	AA158196
8	38.8	4.3	1238	19	AAV48217
9	38.8	4.3	2793	21	AAV48217
10	38.8	4.3	2793	21	AAV48217
11	37.8	4.2	1648	13	AAO25532

C 12	37.2	4.1	1887	22	AAH25568
C 13	37.2	4.1	8269	22	AA164010
C 14	37	4.1	1341	20	AAV98114
C 15	37	4.1	42235	20	AAV98035
C 16	36.2	4.0	1308	18	AAV60756
C 17	36.2	4.0	1308	19	AAV48215
C 18	36.2	4.0	1308	21	AAV251042
C 19	35.4	3.9	690	20	AAV98113
C 20	35.2	3.9	7702	21	AAV88739
C 21	35.2	3.9	7702	21	AAV291908
C 22	35.2	3.9	7702	21	AAV291908
C 23	35.2	3.9	7705	22	AAV98405
C 24	35.2	3.9	7705	22	AAV22684
C 25	35.2	3.9	7741	22	AAV22684
C 26	35	3.9	475	22	AA139079
C 27	34.6	3.9	213	22	AA152138
C 28	34.2	3.8	4134	21	AAV28152
C 29	34	3.8	1307	21	AAV63927
C 30	33.8	3.8	242	22	AAV30127
C 31	33.8	3.8	1032	19	AAV50486
C 32	33.8	3.8	6253	19	AAV40887
C 33	33.8	3.8	6254	19	AAV40887
C 34	33.8	3.8	6254	22	AAV98452
C 35	33.8	3.8	7193	19	AAV50431
C 36	33.8	3.8	9436	15	AAV63499
C 37	33.4	3.7	249	22	AA119369
C 38	33.4	3.7	249	22	AA144559
C 39	33.4	3.7	249	22	AA105056
C 40	33.4	3.7	456	22	AA114406
C 41	33.4	3.7	456	22	AA135779
C 42	33.4	3.7	456	22	AA104229
C 43	33.4	3.7	476	22	AA110073
C 44	33.4	3.7	476	22	AA131321
C 45	33.4	3.7	476	22	AA100016

## ALIGNMENTS

RESULT 1  
AAV31986 standard: cDNA: 898 BP.  
ID AAV31986  
XX AAV31986;  
XX 28-SEP-1998 (first entry)  
XX Mouse Down syndrome-cell adhesion molecule DS-CAM mid cDNA.  
DE DS-CAM; Down syndrome-cell adhesion molecule; neural cell;  
XX signal transduction; trisomy 21; mental retardation;  
KW holoprosencephaly; corpus callosum agenesis;  
KW schizencephaly; diagnosis; assay; mouse; ds; ss.  
XX  
XX Mus sp.  
XX WO9817795-A1.  
XX 30-APR-1998.  
XX 23-OCT-1997; 97WO-US19547.  
XX 25-OCT-1996; 96US-0029322.  
XX  
XX Korenberg JR;  
XX WPI: 1998-271791/24.  
XX New isolated Down's Syndrome-cell adhesion molecule - used to  
XX develop products for detection, diagnosis and therapy of  
XX developmental and neurological abnormalities  
PT

Nucleotide sequenc  
Human polynucleoti  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Rat aortic prefere  
Rat aortic prefere  
Nucleotide sequenc  
Human protein tyro  
Human tyrosine phosph  
Human EST-derived  
Human cDNA encodin  
Human cDNA encodin  
Probe #7765 used t  
Probe #70824 used  
Human NT-CAM gene  
DNA encoding prote  
Mouse sodium chann  
Streptomyces clavu  
Nucleotide sequenc  
Coding sequence of  
Human cDNA clone C  
Streptomyces clavu  
Blood transactibl  
Probe #9302 for ge  
Probe #13245 used  
Probe #5087 used t  
Probe #4339 for ge  
Probe #4465 used t  
Probe #4220 used t  
Probe #6 for gene  
Probe #7 used to m  
Probe #7 used to m









[illegible]

```

XX ds: human; aortic preferentially expressed protein 1; smooth muscle;
KM cell proliferation; developmental stage; tissue plasminogen activator;
XX p21 cell cycle; nitric oxide synthetase; gamma-interferon.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 88..429
FT /tag="a" "Aortic-preferentially-expressed protein-1"
FT /product="
FT /note="No start codon given"
XX
XX W09835040-A2.
XX
XX 13-AUG-1998.
XX
XX 06-FEB-1998; 98WO-US02441.
XX
XX 06-FEB-1997; 97US-0795868.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Hsieh C, Lee M.
XX
XX WPI: 1998-447237/38.
XX
XX P-PSDB; AAW77047.
XX
XX Novel human, rat or mouse aorta or striated-muscle preferentially
XX expressed proteins - useful for treating e.g. atherosclerosis
XX
XX Disclosure: Fig 16; 88pp; English.
XX
XX The aortic preferentially expressed protein 1 (APEG-1) can be used to
XX derive an enhancer/promoter. This linked to a polypeptide coding sequence
XX which regulates smooth muscle cell-specific expression of the polypeptide
XX coding sequence can be used as a method of inhibiting vascular smooth
XX muscle cell proliferation. The nucleic acids are used to direct
XX developmental stage-specific expression of a heterologous polypeptide
XX which is especially selected from tissue plasminogen activator (tPA),
XX p21 cell cycle inhibitor, nitric oxide synthetase, gamma-interferon,
XX atrial natriuretic proteins. These are used to inhibit the proliferation
XX of smooth muscle cells, e.g. for the treatment of atherosclerosis.
XX
XX Sequence 1238 BP; 251 A; 365 C; 387 G; 232 T; 3 other;
XX
XX Query Match 4.18; Score 38.8; DB 19; Length 1238;
XX Best Local Similarity 62.2%; Pred. No. 0.056;
XX Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
XX
XX OY 592 cccgcatgaatgtagatgacgagagagaacccggaatctctcactgttgca 651
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 388 CACCATCTACTGATGACCGCTTTGCAAGTGTAGAAACCATCGCCACGCTCGACGCCA 329
XX
XX OY 652 aaatctgagcgtagatatacactctctccaccctcc 689
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 328 GGATCCGACGCCGACGACGCCACCTCAGCTCCTCTCC 291
XX
XX RESULT 9
XX AAV48218/C
XX ID AAV48218 standard; DNA: 2793 BP.
XX
XX AC AAV48218;
XX
XX 09-NOV-1998 (first entry)
XX
XX Human striated muscle preferentially expressed gene.
XX
XX ds: human; striated muscle preferentially expressed protein;
XX smooth muscle; cell proliferation; developmental stage; gamma-interferon;
XX tissue plasminogen activator; p21 cell cycle; nitric oxide synthetase.

```

```

XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 3..1986
FT /tag="a" "Striated muscle preferentially expressed
FT /product="
FT /note="No start codon given"
XX
XX W09835040-A2.
XX
XX 13-AUG-1998.
XX
XX 06-FEB-1998; 98WO-US02441.
XX
XX 06-FEB-1997; 97US-0795868.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Hsieh C, Lee M.
XX
XX WPI: 1998-447237/38.
XX
XX P-PSDB; AAW77048.
XX
XX Novel human, rat or mouse aorta or striated-muscle preferentially
XX expressed proteins - useful for treating e.g. atherosclerosis
XX
XX Claim 14; Fig 20; 88pp; English.
XX
XX The striated muscle preferentially expressed protein (SPEG) can be linked
XX to an enhancer/promoter derived from an aortic preferentially expressed
XX gene to regulate smooth muscle cell-specific expression. This can
XX be used as a method of inhibiting vascular smooth muscle cell
XX proliferation. The nucleic acids are used to direct developmental
XX stage-specific expression of a heterologous polypeptide which is
XX especially selected from tissue plasminogen activator (tPA), p21 cell
XX cycle inhibitor, nitric oxide synthetase, gamma-interferon, atrial
XX natriuretic proteins. These are used to inhibit the proliferation of
XX smooth muscle cells, e.g. for the treatment of atherosclerosis.
XX
XX Sequence 2793 BP; 499 A; 963 C; 920 G; 411 T; 0 other;
XX
XX Query Match 4.18; Score 38.8; DB 19; Length 2793;
XX Best Local Similarity 62.2%; Pred. No. 0.085;
XX Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
XX
XX OY 592 cccgcatgaatgtagatgacgagagagaacccggaatctctcactgttgca 651
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 1947 CACCATCTACTGATGACCGCTTTGCAAGTGTAGAAACCATCGCCACGCTCGACGCCA 1888
XX
XX OY 652 aaatctgagcgtagatatacactctctccaccctcc 689
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 1887 GGATCCGACGCCGACGACGCCACCTCAGCTCCTCTCC 1850
XX
XX RESULT 10
XX AA251044/C
XX ID AA251044 standard; cDNA: 2793 BP.
XX
XX AC AA251044;
XX
XX 05-JUN-2000 (first entry)
XX
XX Human striated muscle preferentially expressed gene partial cDNA.
XX
XX Human: striated muscle preferentially expressed gene; SPEG;
XX aortic-preferentially-expressed gene-1; APEG-1; striated muscle cell;
XX aorta; smooth muscle cell; antiatherosclerotic; vasotrophic;
XX treatment; diagnosis; vascular disease; atherosclerosis; restenosis;
XX chromosome 2p33-34; ss.
XX

```



```

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 3..1988
FT /*tag- a
FT /product- "SPEG protein"
FT /partial
XX
XX WO200009689-A2.
XX
XX 24-FEB-2000.
XX
XX 11-MAY-1999; 99MO-US10298.
XX
XX 14-AUG-1998; 98US-0134250.
XX 30-APR-1999; 99US-0303069.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Lee M, Hsieh C;
XX
XX MPI: 2000-224334/19.
XX P-PSDB: AAV70078.
XX
XX New gene useful for treating and diagnosing vascular diseases comprises
XX a single gene encoding aortic-specific and striated-specific muscle
XX cell isoforms.
XX
XX Claim 29: Fig 20; 88pp; English.
XX
XX The present sequence is a partial cDNA encoding human striated muscle
XX preferentially expressed gene (SPEG) protein. This protein
XX was detected in striated muscles of skeletal and cardiac tissues.
XX It is one of the two muscle cell protein isoforms encoded by
XX the aortic-preferentially-expressed gene-1 (APEG-1) located on human
XX chromosome 2q33-34. The other protein isoform is specific to aortic
XX smooth muscle cells designated APEG-1 protein. The present sequence can
XX be used to identify striated muscle cell specific promoter which may be
XX useful to direct gene expression in striated muscle cells to treat
XX diseases associated with these muscles. APEG-1 gene is useful for
XX treating and diagnosing vascular diseases such as atherosclerosis and
XX restenosis.
XX
XX Sequence 2793 BP; 499 A; 963 C; 920 G; 411 T; 0 other:
XX
XX Query Match 4.3%; Score 38.8; DB 21; Length 2793;
XX Best Local Similarity 62.2%; Pred. No. 0.085;
XX Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
XX
XX 592 ccccgatgaatgagatagatgagagagaagaacacgaatcctcctcactggttgcga 651
XX 1111111111111111111111111111111111111111111111111111111
XX Db 1947 CACCACTACTCAATGTGACCGCTTTGCAAGTGTAGAAACAGCATGCCACCTCTCGACGCA 1888
XX
XX 652 aaatcgtgagatagatatacctcctccacaccc 689
XX 1111111111111111111111111111111111111111111111111111111
XX Db 1887 GGATTCGCGAGCGCGACACAGCCACCTCAGCGCTCTCC 1850
XX
XX RESULT 11
XX AAQ25532/C
XX ID AAQ25532 standard; DNA; 1648 BP.
XX
XX AAQ25532:
XX
XX 22-MAY-1992 (first entry)
XX
XX Sequence of genomic clone contg. the entire histidine-rich
XX DE protein (HISRP) gene.
XX
XX Malaria vaccine; histidine-rich protein; cytoadherence; ss.
XX KW
XX OS Plasmodium lophurae.

```

```

XX
XX Key Location/Qualifiers
FH sig_peptide 219..359
FT /*tag- a
FT 491..562
FT /*tag- b
FT /product- propeptide
FT Intron 360..490
FT /*tag- c
FT 60..64
FT /*tag- d
FT /label- splice sequence
FT misc_feature 485..490
FT /*tag- c
FT /label- splice sequence
FT mat_peptide 563..1477
FT /*tag- l
XX
XX US116965-A.
XX
XX 25-MAY-1992.
XX
XX 26-AUG-1986; 86US-0900401.
XX PF
XX 26-AUG-1986; 86US-0900401.
XX PR
XX (SILOK ) SILOAN KETTERING INST CANCER.
XX
XX Pologe L, Ravelch JV;
XX
XX MPI: 1992-199590/24.
XX P-PSDB: AAR24393.
XX
XX Histidine-rich protein associated with plasmodium knob phenotype
XX and DNA encoding it, used for in vitro diagnosis of P.
XX falciparum infection.
XX
XX Disclosure: fig 7A-B; 27..35 English.
XX
XX Two variants of HISRP are produced by P. falciparum. One is
XX associated with what is referred to as "knobby" phenotype* (K30) and
XX "knobless phenotype" (K-). The "knobby" and "knobless" phenotypes
XX have been implicated in cytoadherence, which is characteristic of
XX erythrocyte infection. It has now been found that cDNA expressing
XX both K+ and K- HISRP can be obtained by the use of P. lophurae HISRP
XX expressing DNA. The genomic clone encoding sequence from the
XX CC exons, separating the signal peptide encoding sequence from the
XX CC pro-sequence, containing that synthesised to the signal
XX CC the propeptide. Of this, produces synthesised to the signal
XX CC the P. lophurae genome. The sequence of mature proteins is arranged
XX in numerous tandem repeats with up to nine histidine residues in a
XX row, similar to other plasmodium proteins for which sequence data
XX have so far been reported.
XX
XX Sequence 1648 BP; 569 A; 488 C; 150 G; 441 T; 0 other:
XX
XX Query Match 4.2%; Score 37.8; DB 13; Length 1648;
XX Best Local Similarity 58.4%; Pred. No. 0.13;
XX Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
XX
XX 267 ccgcgcatgagatgacatcgtgctgagctcactcctgccaatccggtcttgagctac 326
XX 1111111111111111111111111111111111111111111111111111111
XX Db 1390 TGCATGTCATGATGATGATGTCGTGATGTCATGTCGTGATGTCGTGTCG 1331
XX
XX 327 atcgagatgctgagtgagagagagtgatgatalcaatgatgctgagcaggt 379
XX 1111111111111111111111111111111111111111111111111111111
XX Db 1330 ATCGATGATGATGATGTCGTGTCGTGTCATGTCGTGATGATGATGATGTCGTG 1278
XX
XX RESULT 12
XX AAR25568/C

```







Oy 557 ttcttgacactgagat 573  
| | | | | | | |  
Db 21611 TCCACGTCCTGTACGGT 21595

Search completed: March 4, 2002, 05:23:50  
Job time: 17659 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:12:23 : Search time 226.99 seconds  
(without alignments)  
895.974 Million cell updates/sec

Title: US-08-956-991-8

Sequence: 898  
1 tgcgcgcgcgttcgaagcct.....gcgcacacatcgtcgtga 898

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 351'03 seqs, 113238999 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_patents\_MA:\*

1: /cgn2\_6/plodata/2/ina/5a\_COMB.seq:\*  
2: /cgn2\_6/plodata/2/ina/5b\_COMB.seq:\*  
3: /cgn2\_6/plodata/2/ina/6a\_COMB.seq:\*  
4: /cgn2\_6/plodata/2/ina/6b\_COMB.seq:\*  
5: /cgn2\_6/plodata/2/ina/PCRMUS\_COMB.seq:\*  
6: /cgn2\_6/plodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Match	Length	ID	Description
C 1	38.8	4.3	1225	1	US-08-494-577-11 Sequence 11, Appl
C 2	38.8	4.3	1238	2	US-08-795-868-11 Sequence 11, Appl
C 3	38.8	4.3	2793	2	US-08-795-868-13 Sequence 13, Appl
C 4	36.4	4.1	7218	1	US-08-232-463-14 Sequence 14, Appl
C 5	36.2	4.0	342	2	US-08-494-577-2 Sequence 2, Appl
C 6	36.2	4.0	1308	2	US-08-795-868-2 Sequence 1, Appl
C 7	36.2	4.0	1308	2	US-08-494-577-1 Sequence 1, Appl
C 8	36.2	4.0	1308	2	US-08-795-868-1 Sequence 1, Appl
C 9	31.8	3.5	1755	1	US-08-068-395A-2 Sequence 1, Appl
C 10	31.8	3.5	1755	1	US-08-464-365-2 Sequence 1, Appl
C 11	31.4	3.5	1557	3	US-09-043-123-1 Sequence 1, Appl
C 12	31	3.5	1652	1	US-08-324-243-34 Sequence 34, Appl
C 13	31	3.5	1652	1	US-08-532-390-34 Sequence 34, Appl
C 14	31	3.5	1632	3	US-08-717-294-34 Sequence 34, Appl
C 15	31	3.5	1632	3	US-08-532-390-34 Sequence 34, Appl
C 16	31	3.5	1632	3	US-08-717-294-34 Sequence 34, Appl
C 17	31	3.5	2481	1	US-08-324-243-35 Sequence 35, Appl
C 18	31	3.5	2481	1	US-08-532-390-35 Sequence 35, Appl
C 19	31	3.5	2481	1	US-08-717-294-35 Sequence 35, Appl
C 20	31	3.5	2481	1	US-08-532-390-35 Sequence 35, Appl
C 21	30.8	3.4	2870	2	US-08-468-036-28 Sequence 28, Appl
C 22	30.8	3.4	2870	2	US-08-376-843-28 Sequence 28, Appl
C 23	30.8	3.4	4057	4	US-08-894-997-49 Sequence 49, Appl
C 24	30.6	3.4	3549	4	US-09-008-097-5 Sequence 5, Appl
C 25	30	3.3	3568	4	US-09-103-510B-1 Sequence 1, Appl
C 26	29.8	3.3	375	3	US-09-167-354-8 Sequence 8, Appl
C 27	29.8	3.3	1335	2	US-08-965-090-3 Sequence 3, Appl

C 28	29.8	3.3	1335	3	US-09-165-543-3 Sequence 3, Appl
C 29	29.8	3.3	1335	3	US-09-167-354-6 Sequence 6, Appl
C 30	29.8	3.3	2614	2	US-08-795-868-15 Sequence 15, Appl
C 31	29.8	3.3	2689	2	US-08-985-090-1 Sequence 1, Appl
C 32	29.8	3.3	2689	3	US-09-165-543-1 Sequence 1, Appl
C 33	29.8	3.3	2699	3	US-09-167-354-5 Sequence 5, Appl
C 34	29.6	3.3	206	4	US-09-476-124-11 Sequence 11, Appl
C 35	29.6	3.3	206	4	US-09-476-124-43 Sequence 43, Appl
C 36	29.6	3.3	1013	4	US-09-475-316A-30 Sequence 30, Appl
C 37	29.6	3.3	2424	4	US-09-234-393-1 Sequence 39, Appl
C 38	29.6	3.3	2525	4	US-09-234-393-39 Sequence 39, Appl
C 39	29.6	3.3	2528	4	US-09-234-393-37 Sequence 37, Appl
C 40	29.6	3.3	2528	4	US-09-234-393-41 Sequence 41, Appl
C 41	29.6	3.3	2571	4	US-09-234-393-12 Sequence 12, Appl
C 42	29.6	3.3	2774	4	US-09-363-199B-5 Sequence 5, Appl
C 43	29.6	3.3	4285	4	US-09-040-774-1 Sequence 1, Appl
C 44	29.4	3.3	890	4	US-09-475-316A-34 Sequence 34, Appl
C 45	29.4	3.3	913	4	US-09-475-316A-32 Sequence 32, Appl

## ALIGNMENTS

RESULT 1  
US-08-494-577-11/c  
Sequence 11, Application US/08494577  
Patent No. 5786171  
GENERAL INFORMATION:  
APPLICANT: Lee, Mn-Bn  
TITLE OF INVENTION: Hsieh, Chung-Ming  
TITLE OF INVENTION: ADRIATIC PREFERENTIALLY EXPRESSED GENE AND  
NUMBER OF INVENTIONS: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER RELEASABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,577  
FILING DATE: 22 JUN 1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05434/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1225 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-494-577-11

Query Match 4.3% Score 38.8; DB 1; Length 1225;  
Best Local Similarity 62.2% Pred. No. 0.0056;  
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

0y 592 cccgataatgaaatgacatgacgagagaaacggagatcttctcraactgttgca 651  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



ADDRESS: Fish & Richardson P.C.  
STREET: 225 Richardson Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER REAMABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,577  
FILING DATE: 22-JUN-1995  
CLASSIFICATION: 416  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K  
REGISTRATION NUMBER: 34,819  
REFERENCE/NOTICE NUMBER: 05433/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

```

Query Match          4.08: Score 16.2: DB 1: Length 342:
Best Local Similarity 60.88: Pred. No. 0.019:
Matches 59: conservative 0: Mismatches 38: Indels 0: Gaps 0:

Qy 592 gccgataaactatgaataytgcgagagaagaacccggaattctctcgaactgttgcga 651
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ccccatatttcgttgcacccgttgcacagtgtaagaaaccagcatctcccttcacacacca 242

Qy 652 aaacttcacagcatagatatcactctctctccacactc 688
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ggatcccttcaatccatcagaaatccacgttcgactc 205

RESULT 6
US-08-795-868-2/-
: Sequence 2, Application US/08795868
: Patent No. 5846773
: GENERAL INFORMATION:
: APPLICANT: lee, Mi-Su
: APPLICANT: Hsieh, Chung-Ming
: TITLE OF INVENTION: A SINGLE GENE ENCODING AORTIC-SPECIFIC
: TITLE OF INVENTION: AND STRIATED-SPECIFIC MUSCLE CELL ISOFORMS AND USES THEREOF
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: US
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION NUMBER: US/08/795,868
: FILING DATE: 06-FEB-1997
: CLASSIFICATION: 424
: PRIORITY APPLICATION DATA:

```



APPLICATION NUMBER: 08/494,577  
FILING DATE: 22-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/032001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-795-868-2

Query Match 4.0%; Score 36.2; DB 2; Length 342;  
Best Local Similarity 60.8%; Pred. No. 0.019; Mismatches 18; Indels 0; Gaps 0;  
Matches 59; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 592 ccccgatgaatgatacagcagcagagaagaacgga tctctcctcactgttgcga 651  
DB 301 ccccatattcgttgaccgcttgacagctgagaaacagcatctccctctcagcagcca 242

QY 652 aaatctgacgtagatatactctcctccacacc 688  
DB 241 ccatctcgaacccgacagaccacccctcggcctcctc 205

RESULT 7  
US-08-494-577-1/c  
Sequence 1, Application US/08494577  
Patent No. 5786171  
GENERAL INFORMATION:  
APPLICANT: Lee, Mu-En  
APPLICANT: Hsieh, Chung-Ming  
TITLE OF INVENTION: AORTIC PREFERENTIALLY EXPRESSED GENE AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,577  
FILING DATE: 22-JUN-1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1308 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

US-08-494-577-1

Query Match 4.0%; Score 36.2; DB 1; Length 1308;  
Best Local Similarity 60.8%; Pred. No. 0.043; Mismatches 38; Indels 0; Gaps 0;  
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 592 ccccgatgaatgatacagcagcagagaagaacggaatctctcctcactgttgcga 651  
DB 470 ccccatattcgttgaccgcttgacagctgagaaacagcatctccctctcagcagcca 411

QY 652 aaatctgacgtagatatactctcctccacacc 688  
DB 410 ccatctcgaacccgacagaccacccctcggcctcctc 374

RESULT 8  
US-08-795-868-1/c  
Sequence 1, Application US/08795868  
Patent No. 5846773  
GENERAL INFORMATION:  
APPLICANT: Lee, Mu-En  
APPLICANT: Hsieh, Chung-Ming  
TITLE OF INVENTION: A SINGLE GENE ENCODING AORTIC-SPECIFIC  
TITLE OF INVENTION: AND RELATED-SPECIFIC MUSCLE CELL ISOFORMS AND USES THEREOF  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,868  
FILING DATE: 06-FEB-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/494,577  
FILING DATE: 22-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1308 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-795-868-1

Query Match 4.0%; Score 36.2; DB 2; Length 1308;  
Best Local Similarity 60.8%; Pred. No. 0.043; Mismatches 38; Indels 0; Gaps 0;  
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 592 ccccgatgaatgatacagcagcagagaagaacggaatctctcctcactgttgcga 651  
DB 470 ccccatattcgttgaccgcttgacagctgagaaacagcatctccctctcagcagcca 411

QY 652 aaatctgacgtagatatactctcctccacacc 688  
DB 410 ccatctcgaacccgacagaccacccctcggcctcctc 374

[illegible]

```
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Humicola insolens
FEATURE:
NAME/KEY: CDS
LOCATION: 110..1624
FEATURE: sig.peptide
NAME/KEY: 110..169
LOCATION: 110..169
FEATURE:
NAME/KEY: mal.peptide
LOCATION: 170..1624
US-08-464-365-2

Query Match 3.5%: Score 31.8; DB 1; Length 1755;
Best Local Similarity 56.1%: Pred. No. 1.5;
Matches 60: Conservative 0; Mismatches 47; Indels 0; Gaps 0;

261 gctcatccgcgtgagtgatctgtctgtggtcctccttgcacatccttgcacatcgttctg 320
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1098 GCCTTGATCGCTCGAAGGTGATCTCTCTGATCGAAGGGAAGGACTTCTGTTCTTG 1039

QY 321 gctcatccgcgtgagtgatctgtctgtggtcctccttgcacatccttgcacatcgttctg 367
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1038 GCGACCTCTGATGCGAAGGCGGGAAGGACTTCTGTTCTGAGTT 992

RESULT 11
US-09-043-123-1
Sequence 1, Application US/09043123A
Patent No. 6096521
GENERAL INFORMATION:
APPLICANT: HAAS, Rainer
APPLICANT: ODENBREIT, Stefan
APPLICANT: MEYER, Thomas F.
APPLICANT: BLUM, Andre
APPLICANT: CORFESY-THEULAZ, Irene
TITLE OF INVENTION: NEW ADHESIN FROM HELICOBACTER PYLORI
FILE REFERENCE: 05648004
CURRENT FILING DATE: 1998-06-26
EARLIER APPLICATION NUMBER: DE/195 35 321.8
EARLIER FILING DATE: 1995-09-22
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1557
LENGTH: 1557
TYPE: DNA
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1554)
US-09-043-123-1

Query Match 3.5%: Score 31.4; DB 3; Length 1557;
Best Local Similarity 50.3%: Pred. No. 1.9;
Matches 77: Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 668 tatcaactctctcccaactcttggtagatagtagtgcacatccttgcacatcgttctg 727
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1126 taagctctctcccaactcttggtagatagtagtgcacatccttgcacatcgttctg 1185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 728 tctctctctctcccaactcttggtagatagtagtgcacatccttgcacatcgttctg 787
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1186 aattgtcactcttggtagatagtagtgcacatccttgcacatcgttctg 1245
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 ctctctctctcccaactcttggtagatagtagtgcacatccttgcacatcgttctg 820
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1246 gtagtagatagtagtgcacatccttgcacatcgttctg 1278

RESULT 12
US-08-324-243-34/C
Sequence 34, Application US/08324243
Patent No. 5786464
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL
PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-INTS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA: US/08/324,243
FILING DATE: 19-SEP-1994
APPLICATION NUMBER: 00786/226001
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, PAUL, T
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/226001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 2 0154
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1632 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: Linear
US-08-324-243-34

Query Match 4.5%: Score 31; DB 1; Length 1632;
Best Local Similarity 54.8%: Pred. No. 2.7;
Matches 64: Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 263 ctcatccgcgtgagtgatctgtctgtggtcctccttgcacatccttgcacatcgttctg 322
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 CTGCTGATCGCTCGAAGGTGATCTCTCTGATCGAAGGGAAGGACTTCTGTTCTTG 496

QY 323 gtacatccgcgtgagtgatctgtctgtggtcctccttgcacatccttgcacatcgttctg 381
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 GTTGGTGTCTGATGCGAAGGCGGGAAGGACTTCTGTTCTGAGTT 437

RESULT 13
US-08-532-390-34/C
Sequence 34, Application US/08532390
Patent No. 5795737
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
APPLICANT: HAAS, JURGEN
TITLE OF INVENTION: High Level Expression of Proteins
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:54 ; Search time 6206.58 Seconds  
(without alignments)  
867,408 Million cell updates/sec

Title: US-08-956-991-9\_COPY-500\_1000

Perfect score: 501  
Sequence: 1 gcttgcgcagcatcaata.....aacacagcacttcagatac 501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

1 number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hcc:\*  
10: qb\_est1:\*  
11: qb\_est2:\*  
12: qb\_hcc:\*  
13: qb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	357.2	71.3	744	11	BF348620
2	261.2	52.1	728	11	BE721323
3	139.6	27.9	712	11	BE528580
4	86.4	17.2	1008	13	CNS020JS
5	82.8	16.5	310	11	F13426
6	80.2	16.0	294	11	Z45894
7	53	10.6	612	13	FR0023054
8	46.4	9.3	619	13	FR0030764
9	45.4	9.1	601	13	FR0022062
10	40	8.0	614	13	FR0030753
11	39.8	7.9	412	10	AI945592
12	39.8	7.9	617	11	BE976166

C 13	39.8	7.9	631	11	BE977242	BE977242 bs61d09.y
C 14	39.8	7.9	631	11	BE486312	BE486312 AT20576.5
C 15	39.8	7.9	646	11	BE976689	BE976689 bs54f11.y
C 16	39.8	7.9	822	11	BE486470	BE486470 AT20777.5
C 17	38.4	7.7	428	10	AI596938	AI596938 v148b08.y
C 18	37.8	7.5	359	10	AM481004	AM481004 34091 MAR
C 19	37.8	7.5	1097	13	CNS00815	AL051905 Drosoph11
C 20	37.8	7.4	608	11	BI082263	BI082263 602877444
C 21	36.8	7.3	249	11	R86463	R86463 RABEST032T
C 22	36.8	7.3	634	11	BI365325	BI365325 RES0593.5
C 23	36.8	7.3	635	11	BI237030	BI237030 RE33174.5
C 24	36.6	7.3	454	10	AI756515	AI756515 ELESSTea23
C 25	36.6	7.3	636	10	BE574841	BE574841 F03-P1.T1
C 26	36.4	7.3	426	10	AA400616	AA400616 zu70C08.r
C 27	36.2	7.2	313	10	AI944477	AI944477 bs01h04.y
C 28	36.2	7.2	456	11	R1437	R1437 v170e10.r1
C 29	36.2	7.2	720	13	CNS023MG	AI17965 Tetradon
C 30	36	7.2	260	13	A2563606	A2563606 RPT-23-2
C 31	35.8	7.1	643	10	AA539574	AA539574 LD17768.5
C 32	35.8	7.1	648	11	BI214036	BI214036 RE19729.5
C 33	35.8	7.1	649	11	BI163320	BI163320 RE02690.5
C 34	35.8	7.1	655	11	BI167237	BI167237 RE07573.5
C 35	35.8	7.1	721	11	BI369358	BI369358 RES5515.5
C 36	35.8	7.1	1179	13	CNS040MR	AL269100 Tetradon
C 37	35.6	7.1	699	10	BE662781	BE662781 EST00533
C 38	35.6	7.1	704	10	BE662772	BE662772 EST00523
C 39	35.6	7.1	829	13	BH059023	BH059023 RPT-24-3
C 40	35.4	7.1	882	13	AF029541	AF029541 AF029541
C 41	35.4	7.1	1101	13	CNS016OW	AL107018 Drosoph11
C 42	35.2	7.0	425	10	AM208236	AM208236 M110899
C 43	35.2	7.0	475	10	AL381017	AL381017 M1856A12
C 44	35.2	7.0	500	13	A2340922	A2340922 IM0073M05
C 45	35.2	7.0	535	11	BF004342	BF004342 EST432840

#### ALIGNMENTS

RESULT 1  
BF348620  
LOCUS 602019924F1 NCL CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4155508  
DEFINITION 5', mRNA sequence.  
ACCESSION BF348620  
VERSION BF348620.1 GI:11296215  
KEYWORDS EST.  
SOURCE Human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM9426 row: p column: 05  
High quality sequence stop: 697.  
Location/Qualifiers  
1..744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCL CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"

#### FEATURES

source  
1..744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCL CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"

```

/lab.host="DH10B (T1 phage-resistant)"
/notes="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. primer:Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI-CCAP Library."
BASE COUNT      188 a      196 c      210 g      150 t
ORIGIN

```

	Query Match	71.3%	Score 357.2; DB 11;	Length 744;	
	Best Local Similarity	87.3%;	Pred. No. 2.3e-89;		
Matches	438:	Conservative	0;	Mismatches	56;
				Indels	6;
				Gaps	4;
OY	1 gctllgcacagcatcaatlaaccaccccgagtgaaggcgtgataactgatggcttgaaatgacyg	60			
Dd					
	41 GCTGTTTGGCCAGCATCAACACCACACGCCGTGCAGCGTCGAATCATTGTGGCTGAATGATG	100			
OY	61 cggcgltccaalracctattcaccttccltgaatatagaayacctttggaccacagcttgtac	120			
	101 CGGCGTGCCCATTACACTCTTCTCACACTAGAGTAGAGGCTTTGGACTVCAAG--TTGGAC	158			
OY	121 cacagctcaagcgagacctccccttlccaaaalcctaacaactlctglatalyaccttgaaagaa	180			
Dd					
	159 CACAGCTCAAGAAGACCTCTCTCCAAAGTCH - ACATCCTGTAATGACCTGVAGAAAGCCA	217			
OY	181 cgtagtatgaactcagaatlgtaagatggtgcaaaagcgccgctgctgtgcgataagaagcca	240			
Dd					
	218 CCTGTATGACCTCGAAGATGGCGGCTCAACAAGTGGCGGCTGCGCGAGAAGVAGGCCA	277			
OY	241 acttgcacagctgaactacgaatgacgtacacacacccctcrtcartaaltaatlc. " lglucc	300			
Dd					
	278 ACTTGCTTACGCTGAGACTACGATGGCACTACATAITTCCTCCACTCATTAAGTCAGTTTTCT-	336			
OY	301 acaaagcgaaagaaggctgtgacaaccaaacgaagaaggctlcaaalatcctcgtlyaacatctcty	360			
Dd					
	337 -CAAACGCAAGAGGCGCTGACGACCAACAGAGGGCGCTCAAGATGCTGGTACCATCTCCG	395			
OY	361 catctgtgtcgaggtgtctactgctcttctgtctctgtctgtgctgtgtgtgtgtgtgtgtgtgt	420			
Dd					
	396 TATCTTGATGGGGGCTTGTCTCTCTTGTGTCTCTCTCTGTGGTGGCGAGGAGCGCGG	455			
OY	421 agaacagagagcctgaagaqcctgaagaatgacaaagatctaatgcttgaaa-"lgtccalgagca	479			
Dd					
	456 GGAGCGAGAGGCTGAAGAGGCTGCGAGATGCAAGAAGTTTAACTGMAAGTCCATGATGATGA	515			
OY	480 aaacacacagqacctcaqatac	501			
Dd					
	516 AGAATACCCGCGACTTCAGATTAC	537			
(					
RESULT 2					
BE721323					
LOCUS	BE721323	528 bp	mRNA	EST	25-Apr-2001
DEFINITION	188457 MARC 4BOV Bos taurus CDNA 5', mRNA sequence.				
ACCESSION	BE721323				
VERSION	BE721323.1	GI:10122619			
KEYWORDS	EST.				
SOURCE	COW.				
ORGANISM	Bos taurus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos. 1 (bases 1 to 528)					
REFERENCE					
AUTHORS	Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,D., Fahnenkrug,S.C., Bennett, G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chilko-Mckown,C.G., Pettea,G., Holt,I., Karamcheva,S., Liang,F., Quackenbush,J.D., and Keele,J.W.				
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cdna libraries and construction of a gene index for cattle				
JOURNAL	Genome Res. 11 (4), 626-630 (2001)				
MEDLINE	21180013				
COMMENT	Contact: Smith tpL				

USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: [smith@mail.marc.usda.gov](mailto:smith@mail.marc.usda.gov)  
Single pass sequencing. Bases called and alt. trimmed with phred  
v.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.

FORWARD: AGGAACAGCCTATGACCAT  
 BACKWARD: GTTTCACAGTCACGAC  
 Plate: 98 row: A column: 1  
 Seq primer: ATTATAGTGACACTATAC.

FEATURES  
 source  
 Location/Qualifiers  
 1..528  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 4B09"  
 /listing\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
 library made from pooled tissue from day 20 and day 40  
 embryos."

BASE COUNT 142 a 132 c 154 g 100 t  
 ORIGIN

Query Match 52.1%; Score 261.2; DB 10; length 528;  
 Best Local Similarity 89.3%; pred. NO. 1.4e-62;  
 Matches 293; Conservative 0; Mismatches 33; Indels 2; Gaps 1;

OY	174	gaagcctcagtgatagactgaaatgaaagagtgatcaacagcgccggtgtgacgataag	233
Db	2	GAGGGCAGCTGTGTACCAACTATACATGAGGCGTGTCCACACAGCGCGCTGTGCCACAGAG	61
OY	234	caagcctaacctctgcacacgtctgaactacgattggcagacacaaatccctccacatcaatga	293
Db	62	CAGACCAACTTCCGCCAGCGTCACACTATCATGATGCGCACTACATATCCCCCACTATTATATCA	121
OY	294	gtctctgcacaaagcgaagaaatgctgtaacaaacaaagcagcgtcaagaatcctgtgacca	353
Db	122	GTTGTC -CAAAAGCATATGAAGGGCTGACACACACAGCGGCTCAAGATGCTGTGACCA	179
OY	354	tctctctgatactccagtcacggtctctctacgtctcttgctctgctgtctgtgagcgaagga	413
Db	180	TCTCTCTGATATCTGGTGAGGAGCTCTGTATATCTCTGTTGCTCTCTGCTGTGCTGGAGCA	239
OY	414	gaacatcgaagatcaaaagatcgaagagatcgaagagatgcacaaagatllagatgaatgctca	473
Db	240	GCAGGCTGTATACATACAAAGCTGTGAAGTGGCTGTGAATGATGCAAAAGGTTTACTGTGAATCTCA	299
OY	474	tgaatcaaaacacacacgacgtattcagataac	501
Db	300	TGAGTAAATAATACCCGACCTTCACACAC	327

RESULT 3  
 BF528580 LOCUS 712 bp mRNA EST 11-DEC-2000  
 DEFINITION 602043580P1 NC1.CCAP-Brn67 Homo sapiens cDNA clone IMAGE:4181250  
 5', mRNA sequence.  
 ACCESSION BF528580  
 VERSION BF528580.1 GI:11615943  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 712)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cyabbs-remail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI).  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINI at:  
<http://image.jnl.gov>  
 Plate: LIA9493 row: p column: 19  
 High quality sequence stop: 712.

## FEATURES

source

Location/Qualifiers

1..712  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4181250"  
 /clone\_lib="NCI\_CGAP\_Brn67"  
 /lssue\_type="aneuplastic oligodendrogloma with 1p/19q  
 loss"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: brain; Vector: pCMV-SpK6; Site\_1: NCI;  
 Site\_2: Salt; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 180 a 210 c 204 g 118 t  
 ORIGIN

Query Match 27.9%; Score 139.6; DB 11; Length 712;

Best Local Similarity 60.0%; Pred. No. 1.6e-28;

Matches 300; Conservative 0; Mismatches 174; Indels 26; Gaps 3;

## FEATURES

source

Location/Qualifiers

1..1008  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="c17120"  
 /clone\_lib="c17120"  
 /note="Genoscope sequence ID : CUAG167BE10SP1-end ;  
 pin-on 1"

BASE COUNT 229 a 267 c 251 g 255 t 6 others

## ORIGIN

Query Match 17.2%; Score 86.4; DB 13; Length 1008;

Best Local Similarity 60.0%; Pred. No 1.4e-13;

Matches 162; Conservative 1; Mismatches 102; Indels 5; Gaps 1;

## FEATURES

source

Location/Qualifiers

1..1008  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="c17120"  
 /clone\_lib="c17120"  
 /note="Genoscope sequence ID : CUAG167BE10SP1-end ;  
 pin-on 1"

BASE COUNT 229 a 267 c 251 g 255 t 6 others

## ORIGIN

Query Match 17.2%; Score 86.4; DB 13; Length 1008;  
 Best Local Similarity 60.0%; Pred. No 1.4e-13;  
 Matches 162; Conservative 1; Mismatches 102; Indels 5; Gaps 1;

BASE COUNT 229 a 267 c 251 g 255 t 6 others

## ORIGIN

RESULT 4  
 CNS020JUS 1008 bp DNA GSS 15-MAY-2000  
 LOCUS CNS020JUS 1008 bp DNA GSS 15-MAY-2000  
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone  
 167120 of library G from Tetraodon nigroviridis, genome survey

sequence.

## ACCESSION

AL214561

## VERSION

AL214561.1 G1:7873380

## KEYWORDS

GSS: genome survey sequence.

## SOURCE

Tetraodon nigroviridis.

## ORGANISM

Tetraodon nigroviridis

## REFERENCE

1 (bases 1 to 1008)

## AUTHORS

Roest-Crollius, H., Jallion, O., Basilva, C., Fizames, C., Fisher, C.,  
 Bernot, A., Winkler, P., Brotter, P., Quetier, F., Saurin, W., Bernot, A. and  
 Weissbach, J.

## TITLE

Characterization and repeat analysis of the compact genome of the  
 freshwater puterfish Tetraodon nigroviridis

## JOURNAL

Unpublished

## AUTHORS

Roest-Crollius, H., Jallion, O., Basilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Winkler, P., Brotter, P., Quetier, F., Saurin, W. and  
 Weissbach, J.

## TITLE

Human gene number estimate provided by genome wide analysis using  
 Tetraodon nigroviridis DNA sequence

## JOURNAL

Unpublished

## AUTHORS

3 (bases 1 to 1008)

## REFERENCE

Genoscope.

## TITLE

Direct Submission

## JOURNAL

Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

## COMMENT

This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetraodon nigroviridis  
 genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/tetraodon>.

## FEATURES

Location/Qualifiers

1..1008  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="c17120"  
 /clone\_lib="c17120"  
 /note="Genoscope sequence ID : CUAG167BE10SP1-end ;  
 pin-on 1"

## ORIGIN

Query Match 17.2%; Score 86.4; DB 13; Length 1008;  
 Best Local Similarity 60.0%; Pred. No 1.4e-13;  
 Matches 162; Conservative 1; Mismatches 102; Indels 5; Gaps 1;

## ORIGIN

RESULT 5  
 F13426 310 bp mRNA EST 15-MAR-1995  
 LOCUS F13426 310 bp mRNA EST 15-MAR-1995  
 DEFINITION HSC2XN021 normalized infant brain cDNA Homo sapiens cDNA clone



```

c-2xa02, mRNA sequence.
ACCESSION F13426
VERSION F13426.1 GI:710043
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 310)
AUTHORS Auffray,C., Behar,G., Bois,F., Bonchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Hougaite,R., Juneau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Marique-Samson,R., Pictu,G., Pouillot,Y.,
Sebastien Kabakchis,C. and Jossier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
COMMENT Contact: Genethon
Genexpress-genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, Bp60 91002 I.RY Cedex, FRANCE
Tel : 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_idt: C: Genexpress-sequence_idt: y2c-2xa02
Seq primer: (-21)Mt3-universal.
FEATURES
source Location/Qualifiers
1..310
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2xa02"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lambda BA: Site 1: HindIII;
Site 2: NotI; sex=Female; dev-stage=3 months old;
isolate-muscular atrophy patient; tissue-type-total brain"
1: total mRNA was oligo-(dty) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lambda BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
BASE COUNT 79 a 73 c 89 g 66 t 3 others
ORIGIN
1 cy Match 16.5%; Score 82.8; DH 11; Length 310:
2 st Local Similarity 87.1%; Pred. No. 9, 3e-13;
Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.
Oy 386 ttctgcttcctgctgatttgaggagagacgcgcggcgag-acaggctaagaagacctgcga 445
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Nb 2 TTTCGCCCTCCGTGGTTGCGGAGAGAGCGCGCGGAG-ACAGGCTAAGAAGCGCTCCGA 60
Oy 446 gatcgaagaagttagctgaaadtcgatcatgacaagaacacgcgcactcaaatatc 501
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GATCAAGAAGACTTAGCTGAATATCTCATGATGAAGATAACCGGACTTACGNTAC 116
RESULT 6
LOCUS Z45894 294 bp mRNA EST 14-NOV-1994
DEFINITION HSC2HM041 normalized infant brain cDNA Homo sapiens cDNA clone
C-2WM04, mRNA sequence.
ACCESSION Z45894
VERSION Z45894.1 GI:575128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 294)	Aufroy C., Behar G., Bois F., Bouchier C., da Silva C., Devignes M.D., Duprat S., Hougatete R., Juneau M.N., Lamy B., Lorenzo F., Mitchell H., Marique-Samson R., Pietu G., Pouillot Y., Sebastian Kakkilich C. and Tessier A.	IMAGE: molecular integration of the analysis of the human genome and its expression	C. R. Acad. Sci. III, Sect. Vie 318 (2), 263-272 (1995)	
95277534	Contact: Genethon	Genexpress: Genethon	Genethon Centre de recherche sur le Genome Humain 1, rue de l'Internationale, B160 91002 Evry Cedex, FRANCE Tel: 33169472800 Fax: 33160778698 Email: genexpress@genethon.fr Single read. Genexpress library: Idt: C; Genexpress: sequence_idt: ylic-zwh04 Seq primer: (-21)M3.univers	
FEATURES	Source	1..294		
	/organism "Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="c-zwh04"			
	/clone_lib="normalized infant brain cDNA"			
	/sex="Female"			
	/tissue_type="total brain"			
	/dev_stage="3 months old"			
	/note="Organ: brain; Vector: lambda BA; Site: 1; HindIII; Site: 2; NotI; sex: female; dev: stage-3 months old; isolate-muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lambda BA vector. Clone library from B.Souares, Psychiatry Dept. Columbia University, USA. Normalization method: Benlo Soares, p.N.A.S in press"			
BASE COUNT	79 a	66 c	87 g	58 t
ORIGIN			4 others	
Query Match	16.0%;	Score 80.2;	DB 11;	Length 294;
Best Local Similarity	89.7%;	Pred. No. 4.9e-12;		
Matches	96;	Conservative	0;	Mismatches 10;
			Indels	1;
			Gaps	1;
QY	395	ct gct ggt tat g g g a a g a t c g g t a g a t c t g a a g c t g a a g g c t g a g a t g c a a g	454	
Db	2	ct g c t g c t g c t g c t g c a a g a a a c t g a a g a g c t g a a g c t g c c a c a t t g c a a g	60	
QY	455	a g t t a a c t a a a t g c t c a t g a g a a a a a c a c a g g a c t c a g a t a c	501	
Db	61	a g t t a a c t g a a a t g c t c a t g a g a a a a a c a c a g g a c t c a g a t a c	107	
RESULT 7				
LOCUS	FR0022054	612 bp	DNA	GSS
DEFINITION	F. rubripes GSS sequence, clone 070M05aA12, genomic survey sequence			10-DEC-1997
ACCESSION	AL014925			
VERSION	AL014925.1	GI:2681293		
KEYWORDS	GSS: genome survey sequence.			
SOURCE	Takifugu rubripes.			
ORGANISM	Takifugu rubripes			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Takifugu.			
REFERENCE	1 (bases 1 to 612)			
AUTHORS	Elgar G., Clark M., Smith S., Weck S., Warner S., Uranian Y., Williams G. and Brenner S.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-DEC-1997) WRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hmp.mrc.ac.uk			
COMMENT	Vector: pBluescript II KS			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei





```

/sux="male"/
/dev.stage="0-3 day old Ore-R males"
/lab.host="Plates AT_10-AT_120: DH5-alpha. plates
AT_121-AT_319: DH5-alpha."juja"
./note="Organ: ADULT testes; Vector: pOTB7; Site: 1; Exon:1
Site_2: Xho1; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
157 d 202 c 167 g 105 l

```

[illegible]

RESULT	15			
HE976689/c				
LOCUS	BE976689	646 bp	mRNA	EST 04-OCT-2000
DEFINITION	bs24111.y1 Drosophila melanogaster adult testis library Drosophila melanogaster cDNA clone bs24111 5', mRNA sequence.			

ACCESSION	BE9766689
VERSION	BE9766689.1
	GI:10606415

ORGANISM *Drosophila melanogaster*

Muscomorpha; Ephydroides

**TITLE** *Drosophila melanogaster*

COMPLEMENT  
CONTACT: BRIAN STREIB  
Laboratory of Cellular

6 Center Drive MSC 2715, Bldg 6, Km H-13, Beltsville, MD 20892 USA  
Fax: (301) 496 5239  
Email: [oliverh@nih.gov](mailto:oliverh@nih.gov),  
<http://www.niddk.nih.gov/intram/people/boliverer.htm>  
Tissue isolation and library construction performed at the National  
Institute of Diabetes and Digestive and Kidney Diseases, NIH (see  
<http://www.niddk.nih.gov/intram/people/boliverer.htm>). DNA sequencing  
and analyses performed by National Institutes of Health Intramural  
Sequencing Center (NISC; see <http://www.nisc.nih.gov>).  
Plate: 54 row: 1 column: 1  
Seq primer: M13Rpl reverse primer (AatI).

source 1 . . 646

```

1. .646
/organism="Drosophila melanogaster"
/strain="y[*] w[67c] y"
/db_xref="taxon:7227"
/clone="hs54f1"
/clone_lib="Drosophila melanogaster adult testis library"
/sex="male"
/dev_stage="1-5 day adult"
/lab_host="Solk (Stratagene)"
/note="Organ testis: Vector: pBlueScript SK (Stratagene); Site:1: EcoR1; Site:2: Xho1; Testes dissected from 1-5 day adult y[*] w[67c] y males raised at 25°C. RNA isolated using Trizol (Life Technologies) and a single round of Poly(A) selection using Oligotex (Qiagen). cDNA

```

BASE COUNT  
ORIGIN

184 a 143 c 151 g 166 t

library constructed using Stratagene ZAP-cDNA synthesis kit. Oligo dt-primed, size fractionated -1.6 kb, and directionally cloned at EcoRI and XhoI in Uni-ZAP XR. Following a single round of amplification plus-script plasmids were mass excised. A distribution channel of clones is being sought, but not currently available. Requests for clones cannot be honored."

[illegible]

Search completed: March 4, 2002, 02:15:05  
Job time: 6510 sec

Search completed: March 4, 2002, 02:15:05  
Job time: 6510 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:23:50 ; Search time 578.64 Seconds  
(without alignments)  
742.292 Million cell updates/sec

Title: US-08-956-991-9\_COPY\_500\_1000

Perfect score: 501  
Sequence: 1 gcttccgcgcagcatcaata.....aacacacgactcagatcac 501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

{ number of hits satisfying chosen parameters: 1861242

Minimum nb seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N.Geneseq\_1101.\*  
2: /SID2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	2173	19	AAV31987	Mouse Down syndrom
2	390.2	77.9	6604	19	AAV31981	Human Down syndrom
3	304.2	60.7	388	19	AAV31982	Human Down syndrom
4	179	35.7	6413	19	AAV31988	Human Down syndrom
5	140.8	28.1	5824	22	AA158196	Human Down syndrom
6	35.6	7.1	1627	21	AAV31995	Human Down syndrom
7	35.6	7.1	1627	21	AAV31995	Human Down syndrom
8	33.8	6.7	1171	12	AAV31995	Human Down syndrom
9	33	6.6	692	21	AAV31995	Human Down syndrom
10	32.8	6.5	533	21	AAV31995	Human Down syndrom
11	32.8	6.5	585	19	AAV31997	Cytochrome P45011d

12	32.8	6.5	1494	16	AAO87729	Human auxillary cy
13	32.8	6.5	1494	16	AAO87730	Human auxillary cy
14	32.8	6.5	1494	16	AAO87731	Human auxillary cy
15	32.8	6.5	1494	16	AAO87732	Human auxillary cy
16	32.8	6.5	1494	17	AAO87733	Human auxillary cy
17	32.8	6.5	1494	17	AAO87734	Human auxillary cy
18	32.8	6.5	1494	17	AAO87735	Human auxillary cy
19	32.8	6.5	1494	17	AAO87736	Human auxillary cy
20	32.8	6.5	1494	17	AAO87737	Human auxillary cy
21	32.8	6.5	1494	17	AAO87738	Human auxillary cy
22	32.8	6.5	1494	17	AAO87739	Human auxillary cy
23	32.6	6.5	189	22	AAO87740	Human auxillary cy
24	32.6	6.5	189	22	AAO87741	Human auxillary cy
25	32.6	6.5	189	22	AAO87742	Human auxillary cy
26	32.6	6.5	189	22	AAO87743	Human auxillary cy
27	32.6	6.5	189	22	AAO87744	Human auxillary cy
28	32.6	6.5	189	22	AAO87745	Human auxillary cy
29	32.6	6.5	189	22	AAO87746	Human auxillary cy
30	32.6	6.5	189	22	AAO87747	Human auxillary cy
31	32.6	6.5	189	22	AAO87748	Human auxillary cy
32	32.6	6.5	189	22	AAO87749	Human auxillary cy
33	32.6	6.5	189	22	AAO87750	Human auxillary cy
34	32.6	6.5	189	22	AAO87751	Human auxillary cy
35	32.6	6.5	189	22	AAO87752	Human auxillary cy
36	32.6	6.5	189	22	AAO87753	Human auxillary cy
37	32.6	6.5	189	22	AAO87754	Human auxillary cy
38	32.6	6.5	189	22	AAO87755	Human auxillary cy
39	32.6	6.5	189	22	AAO87756	Human auxillary cy
40	32.6	6.5	189	22	AAO87757	Human auxillary cy
41	31.8	6.3	435	22	AAO87758	Human auxillary cy
42	31.8	6.3	435	22	AAO87759	Human auxillary cy
43	31.6	6.3	435	22	AAO87760	Human auxillary cy
44	31.6	6.3	435	22	AAO87761	Human auxillary cy
45	31.6	6.3	435	22	AAO87762	Human auxillary cy

## ALIGNMENTS

RESULT 1	AAV31987	standard; cDNA: 2173 BP.
ID	AAV31987	
AC	AAV31987	
XX	28-SEP-1998	(first entry)
DT	28-SEP-1998	
XX	Mouse Down syndrome-cell adhesion molecule DS-CAM 3' cDNA.	
DE	DS-CAM: Down syndrome-cell adhesion molecule; neural cell;	
XX	signal transduction; trisomy 21; mental retardation;	
KW	holoprosencephaly; corpus callosum agenesis;	
KW	schizencephaly; diagnosis; assay; mouse; ds; ss.	
XX		
OS	Mus sp.	
XX		
PN	W09817795-A1.	
XX	30-APR-1998.	
PD		
XX	23-OCT-1997; 97MO-US19547.	
PF		
XX	25-OCT-1996; 96US-0029322.	
PR		
XX	(CEDA-) CEDAKS SINAI MEDICAL CENT.	
PA		
XX	Korenberg JR;	
PI	WPI; 1998-271791/24.	
XX		
DR	New isolated Down's Syndrome-cell adhesion molecule - used to	
XX	develop products for detection, diagnosis and therapy of	
PT	developmental and neurological abnormalities	
PT		

XX Claim 2: Page 81-83; 109pp; English.

XX This cDNA sequence comprises the 3' region of a cDNA clone  
XX for murine Down syndrome-cell adhesion molecule (DS-CAM), a  
XX member of a novel subclass of the Ig superfamily with homology to  
XX neural cell adhesion molecules. The middle region (see AAV31981)  
XX and 5' region (see AAV31985) of the clone are also provided. The  
XX murine DS-CAM clone was isolated from a C57 Black/6 mouse brain  
XX cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31988) as  
XX probe. The invention also provides human DS-CAM proteins (see  
XX AAV42086-87), as well as expression vectors and host cells,  
XX transgenic animals, antisense oligonucleotides, and primers useful  
XX for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are  
XX associated with developmental and neurological processes. They can  
XX be used in e.g. neural prosthetic devices used in entubulation  
XX methods of repairing (regenerating) damaged or severed peripheral  
XX nerves. The products can also be used in detection, diagnosis and  
XX therapy of developmental and neurological abnormalities such as  
XX Down syndrome, mental retardation, holoprosencephaly, agenesis of  
XX the corpus callosum, or schizencephaly. Antisense oligonucleotides  
XX are used for inhibiting translation of mRNA.

XX Sequence 2173 BP; 585 A; 598 C; 554 G; 436 T; 0 other;

Query Match 100.0%; Score 501; DB 19; Length 2173;  
Best Local Similarity 100.0%; Pred. No. 5,1e-148;  
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcttttcgcacatcaatcaaccacccaggtgagtgatctgattgctggaatgacg 60  
DB 500 gcttttcgcacatcaatcaaccacccaggtgagtgatctgattgctggaatgacg 559  
QY 61 cggctgacatcaatcaatcaatctctggaatgacaccccttgagacacgcttgac 120  
DB 560 cggctgacatcaatcaatcaatctctggaatgacaccccttgagacacgcttgac 619  
QY 121 caccagctcagcagcctcccttccaaagctcctcaatctctggaatgacacacga 180  
DB 620 caccagctcagcagcctcccttccaaagctcctcaatctctggaatgacacacga 679  
QY 181 cgtggtatgacatcagatgagagtgatgcaacagcgcgcgtgtgagataagcaagca 240  
DB 680 cgtggtatgacatcagatgagagtgatgcaacagcgcgcgtgtgagataagcaagca 739  
QY 241 acttgcacgctggaatcagatgagatgacatccctccatctctggaatgacatgac 300  
DB 740 acttgcacgctggaatcagatgagatgacatccctccatctctggaatgacatgac 360  
QY 301 acaaaaggaagaagggctgacacacgaagagctcaagatccctgacatctcctg 360  
DB 800 acaaaaggaagaagggctgacacacgaagagctcaagatccctgacatctcctg 859  
QY 361 catctgctcggggtctactgctcttctgtctcgtcgtgtgagaggaagagcggc 420  
DB 860 catctgctcggggtctactgctcttctgtctcgtcgtgtgagaggaagagcggc 919  
QY 421 agagcagagctgaaagagctgagatgcaagaagctttagctgaatctctatgacaa 480  
DB 920 agagcagagctgaaagagctgagatgcaagaagctttagctgaatctctatgacaa 979  
QY 481 aaacacacgagctcagatc 501  
DB 980 aaacacacgagctcagatc 1000

RESULT 2  
AAV31981  
ID AAV31981 standard; cDNA; 6604 BP.  
XX  
AC AAV31981;  
XX

DT 28-SEP-1998 (first entry)

XX Human Down syndrome-cell adhesion molecule DS-CAM1 cDNA.

DE DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;

KW signal transduction; trisomy 21; mental retardation;

KW holoprosencephaly; corpus callosum agenesis;

KW schizencephaly; diagnosis; assay; human; ds; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 453..6185

FT sig-peptide 452..521

FT mat-peptide 522..6162

FT /tag= c

PN M09817795-A1.

PD 30-APR-1998.

PF 23-OCT-1997; 97WU-US19547.

PR 25-OCT-1996; 96US-0029322.

XX (CEDA-) CEDARS SINAI MEDICAL CENT.

PA Korenberg JK;

PI WPI: 1998-271791/24.

DR P-PSDB: AAV42086.

XX New isolated Down's Syndrome-cell adhesion molecule - used to

PT develop products for detection, diagnosis and therapy of

PT developmental and neurological abnormalities

XX Claim 3: Page 65-72; 109pp; English.

PS This cDNA clone codes for Down syndrome-cell adhesion molecule

XX DS-CAM1 (see AAV42086), a cell surface glycoprotein belonging to a

CC novel subclass of the Ig superfamily with highest homology to

CC neural cell adhesion molecules. A modified direct cDNA selection

CC technique was applied to bacterial and pl artificial chromosomes

CC between ETS2 and MX1 by using cDNA from trisomy 21 human foetal

CC brain. A unique cDNA fragment, designated E51 (see AAV31982), was

CC obtained and used to screen a trisomy 21 human foetal brain (14-wk

CC gestation) cDNA library. Further clones were obtained by exon

CC trapping. The complete DS-CAM1 cDNA sequence was deduced from

CC overlapping clones. A splice variant cDNA (see AAV31988), encoding

CC non-membrane bound DS-CAM2 (see AAV42087), was also identified. The

CC DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome

CC 21q22.2-22.3. The invention also provides murine DS-CAM partial

CC sequences (see AAV31985-87), expression vectors and host cells,

CC transgenic animals, antisense oligonucleotides, and primers useful

CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are

CC associated with developmental and neurological processes. They can

CC be used in e.g. neural prosthetic devices used in entubulation

CC methods of repairing (regenerating) damaged or severed peripheral

CC nerves. The products can also be used in detection, diagnosis and

CC therapy of developmental and neurological abnormalities such as

CC Down syndrome, mental retardation, holoprosencephaly, agenesis of

CC the corpus callosum, or schizencephaly. Antisense oligonucleotides

CC are used for inhibiting translation of mRNA.

XX Sequence 6604 BP; 1673 A; 1827 C; 1768 G; 1336 T; 0 other;

Query Match 77.9%; Score 390.2; DB 19; Length 6604;  
Best Local Similarity 88.8%; Pred. No. 8.3e-113;  
Matches 445; Conservative 0; Mismatches 53; Indels 3; Gaps 2;

[illegible]

XX This polynucleotide comprises cDNA fragment E51 that was isolated  
CC from a human trisomy 21 foetal brain (14-wk gestation) cDNA  
CC library following a modified direct cDNA selection technique  
CC applied to bacterial and pl artificial chromosomes between ET52  
CC and MX1. E51 was used as a probe to screen the trisomy 21 foetal  
CC brain library. 62 clones were isolated from the 2 million clones  
CC of the original library. overlapping clones were obtained that  
CC were used to deduce a full-length sequence (see AAV31981) coding  
CC for novel down syndrome-cell adhesion molecule DS-CAM (see  
CC AAV42086). A splice variant, DS-CAM2 (see AAV31988), was also  
CC identified. DS-CAM polypeptides are associated with developmental  
CC and neurological processes. The polypeptides and nucleic acids  
CC are used to develop products for the detection, diagnosis and  
CC therapy of developmental and neurological abnormalities.  
XX  
SQ Sequence 388 BP; 83 A; 119 C; 95 G; 91 T; 0 other;

Query Match	60.7%	Score 304.2;	DB 19;	Length 388;
Best Local Similarity	89.5%;	Pred. No. 3.2e-86;		
Matches 350;	Conservative 0;	Mismatches 38;	Indels 3;	Gaps 2.

[illegible]

RESULT	4
AAV31988	
ID	AAV31988 standard; cDNA; 6413 bp.
XX	
AC	AAV31988;
XX	
DT	28-SEP-1998 (first entry)
XX	
DE	Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.
XX	
KW	DS-CAM2; down syndrome-cell adhesion molecule; neural cell;
KW	signal transduction; trisomy 21; mental retardation;
KW	holoprosencephaly; corpus callosum agenesis;
KW	schizencephaly; diagnosis; assay; human; ds; ss.
XX	
OS	Homo sapiens.
XX	
Key	location/qualifiers
FT	453..5168
CDS	/*tag- a
FT	
XX	





OY		2	ccttttcgcagcgtccataatcacaccgcaagtgaagcttgaaactcgattggctcguaaatgatgcgcg	61
Db	4121	ccttccaacccatatcaactcccacgcatgctcygcttaaaccttgacaaggcttgaaaacaatgag	4180	
OY		62	ggctctgtccaaactcaacctccattcaactcttgaataacagaacctttgggaccaacggtcttagcc	121
Db	4181	ggctgcaccttatccacagccatcgtctcttgagaglaacggccaaagygaacccugygcctlyg---	4237	
OY		122	aacacrtcagrcygscctcccccttcccaaagltccctaacaattclqatayacctgcagaagaagccac	181
Db	4238	-cagyygcctccryggccaacacgctccoggyagngtjtlcttgcacygaucrcygagaygcac	4296	
OY		182	tgcglaclyaaactlvcagatgagagtgtylcasaaragaccgagctgtgcgagtaaqaacagccaa	241
Db	4297	tgtgtacagatcgtgcagtagaygccttgcaacagtgccgggtctgcggccaalgsaaacagccaa	4356	
OY		242	ccttcgccagcgtgaacttacgatatggcagtagacaalcctccactcaacttaagtcgaatllytcca	301
Pf	4357	gttcgcacaccccttgcactacgaatgagcagaccac---ttcaaccataagtlctgtccaaag	4413	
OY		302	caaaagcaagaagaygcctgcacaacaaacgaagycctcaaga ccttcgttaarattcvtgc	361
Db	4414	lgaaag-----galysatgtaayaawcgtlctacacalcgctgc	4453	
OY		362	atccctgcgtcgvggtcttaactgtcctcttgttgtcctctgcgtgtgtctgcgaagayaganrgya	421
Db	4454	cctgtcacactlvcgccaacacgcgggytgcactgtccttccalcgtacgcagaagaagagaaq	4513	
OY		422	gagcagagagctgaagagagctgagagatgcaaaagattacgtgaaatgctcatyagcaaa	481
Db	4514	gagaaacggcctgaaagcgatcccgagatgcagaagaagctllgcagaagaattgtgataagcaag	4573	
OY		482	aacatarcgacttcagatalc	501
Db	4574	aacataagaagctttagcac	4593	
RESULT	6			
AAC47995/C				
ID	AAC47995 standard; DNA: 1627 BP.			
XX				
AC	AAC47995;			
XX				
DT	18-OCT-2000 (first entry)			
XX				
DE:	Arabidopsis thaliana DNA fragment SEQ ID NO: 55880.			
XY				
KW	Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.			
OS	Arabidopsis thaliana.			
XX				
PN	EPJ03405-A2.			
PD				
XX	06-SEP-2000.			
PF				
XX	25-FEB-2000; 2000EP-0301439.			
PR				
XX				
PR	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125788.			
PR	25-MAR-1999; 99US-0126285.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127462.			
PR	06-APR-1999; 99US-0128234.			
PR	08-APR-1999; 99US-0128714.			
PR	16-APR-1999; 99US-0129845.			
PR	19-APR-1999; 99US-0130077.			
PR	21-APR-1999; 99US-0130449.			
PR	23-APR-1999; 99US-0130510.			

PR	23-APR-1999;	9905-0130891;
PR	28-APR-1999;	9905-0131449;
PR	30-APR-1999;	9905-0132048;
PR	30-APR-1999;	9905-0132407;
PR	04-MAY-1999;	9905-0132484;
PR	05-MAY-1999;	9905-0132485;
PR	06-MAY-1999;	9905-0132486;
PR	06-MAY-1999;	9905-0132487;
PR	07-MAY-1999;	9905-0132633;
PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134218;
PR	14-MAY-1999;	9905-0134219;
PR	14-MAY-1999;	9905-0134221;
PR	18-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134768;
PR	19-MAY-1999;	9905-0134941;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135373;
PR	24-MAY-1999;	9905-0135629;
PR	25-MAY-1999;	9905-0136021;
PR	27-MAY-1999;	9905-0136392;
PR	28-MAY-1999;	9905-0136782;
PR	01-JUN-1999;	9905-0137222;
PR	03-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137502;
PR	07-JUN-1999;	9905-0137724;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138540;
PR	10-JUN-1999;	9905-0138847;
PR	14-JUN-1999;	9905-0139119;
PR	16-JUN-1999;	9905-0139452;
PR	18-JUN-1999;	9905-0139453;
PR	17-JUN-1999;	9905-0139452;
PR	18-JUN-1999;	9905-0139454;
PR	18-JUN-1999;	9905-0139455;
PR	18-JUN-1999;	9905-0139456;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139460;
PR	18-JUN-1999;	9905-0139461;
PR	18-JUN-1999;	9905-0139462;
PR	18-JUN-1999;	9905-0139463;
PR	18-JUN-1999;	9905-0139750;
PR	18-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139819;
PR	22-JUN-1999;	9905-0140354;
PR	23-JUN-1999;	9905-0140635;
PR	24-JUN-1999;	9905-0140824;
PR	28-JUN-1999;	9905-0140941;
PR	28-JUN-1999;	9905-0141287;
PR	30-JUN-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142505;
PR	06-JUL-1999;	9905-0142590;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0144624;
PR	15-JUL-1999;	9905-0144805;
PR	16-JUL-1999;	9905-0144806;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144332;

PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-015, 84.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161411.
PR	28-OCT-1999;	99US-0161420.
PR	28-OCT-1999;	99US-0161492.
PR	28-OCT-1999;	99US-0161493.
PR	29-OCT-1999;	99US-0162142.

	Query Match	7.1%	Score 45.6;	DB 21;	Length 1627;	
	Best Local Similarity	52.7%;	Pred. No. 0.41;			
	Matches 77;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0	
OY	338 aaatccctcgtgcacattcccgaclcccgctcggggtctactgcttlygtctcag					397
Db	471 AACAGCTTTGACATCATCTCTCCACATGCCAGCGTCGAATAAGCCTTGCTGTGGCTGCGTGG					412
OY	398 ctggcttgctgcggagacgcaqcgagagcagagctcaagaaggcttgagagtcgaagaagt					457
Db	411 CTGTTGATTAGACCCTGATTGCGAAGCGTAGGATGATCTGCTGATTAAGACCTTGATT					352
OY	458 tttagctgaatgccalcagcaaaaa	483				
Db	351 ATACGCAACATATTGCATGACGATACA	326				
RESULT 7						
AAC41871/c						
ID	AAC41871 standard; DNA: 1679 BP.					
XX						
AC	AAC41871:					
XX						
DJ	17-OCT-2000 (first entry)					
DE						
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 34446.					
KM	Hybridisation assay; gene-fc mapping; gene expression control;					
KW	protein identification; signal transduction pathway;					
XX	metabolic pathway; promoter; termination sequence; ss.					
OS	Arabidopsis thaliana.					
XX						
PN	EPI034405-A2.					
XX						
PD	06-SEP-2000.					
XX						
PE	25-FEB-2000; 2000EP-0301439.					
XX						
PR	25-FEB-1999; 990US-0121H25.					
PR	05-MAR-1999; 990US-0123I80.					
PR	09-MAR-1999; 990US-0123548.					
PR	23-MAR-1999; 990US-0125788.					
PR	25-MAR-1999; 990US-0126264.					
PR	29-MAR-1999; 990US-0126785.					
PR	01-APR-1999; 990US-0127462.					
PR	06-APR-1999; 990US-0128234.					
PR	08-APR-1999; 990US-0128714.					
PR	16-APR-1999; 990US-0129845.					
PR	19-APR-1999; 990US-0130077.					
PR	21-APR-1999; 990US-0130449.					

PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132487.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137724.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138099.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145214.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 02-AUG-1999; 99US-0147038.  
PR 03-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149175.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151086.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151940.  
PR 07-SEP-1999; 99US-0152341.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.

Query Match	7.1%	Score 35.6	DB 2	Length 1679
Best Local Similarity	52.7%	pred. No. 0.41	Mismatches 69	Indels 0
Matches	77	Conservative	0	Gaps 0
Oy	338	aagatcctcgtagaccatctctgcatactcgtgctggggtlcttactgctcttggccttcy	397	
Db	523	AAGAGCTCTTGACATCATCTCCACATCCAGCGCATTAACCTCTTGCTGCTCGCTGGTG	464	
Oy	398	ctgattgctgcygaggaagacgycgagagcgagagcgtgaagctgtagagatlycacaagyl	457	
Db	463	CTGTGTGATTAACCCCGATTATGTGAAGCTGACAGATGATGATCTCTGATATAGGACCTTGATT	404	
Oy	458	ttagctgaaatcgtcatcgaacaaa	483	
Db	403	ATACGGAACATATTGATGACGATACA	378	

[illegible]

PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 11-MAY-1999; 9905-0132488.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-013529.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 24-JUN-1999; 9905-0140823.  
PR 28-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142380.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143674.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 23-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147048.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147945.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151970.  
PR 07-SEP-1999; 9905-0152463.  
PR 10-SEP-1999; 9905-0154070.  
PR 13-SEP-1999; 9905-0154078.  
PR 15-SEP-1999; 9905-0154079.  
PR 16-SEP-1999; 9905-0154079.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155149.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.

```
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 22-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 25-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161993.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.
```

```
Query Match 6.6%; Score 33; DB 21; Length 692;
t Local Similarity 60.7%; Pred. No. 1.8;
ches 54; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Oy 373 ggtctactgtctcttctgtctgtctgtgtgagagagagagagagagct 432
|| |||| || ||| | || |||| |||| |||| |||| |||| |||| ||
Db 347 gggcttaatgactcaatcacgatacgtatgttcgaggaactgagcagcagcttga 406
|| || || || || || || || || || || || || || || || || ||
Oy 433 gaagagctgagagatgcaagaagttag 461
| | | | | | | | | | | | | | | | | |
Db 407 gtggttgaagaagctcaagcattcag 435
```

```
RESULT 10
AAC38550
ID AAC38550 standard: DNA: 533 BP.
```

XX AAC38550:

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SKO ID No: 21378.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000: 2000EP-0301439.

XX 25-FEB-1999: 99US-0121825.

PR 05-MAR-1999: 99US-0123180.

PR 09-MAR-1999: 99US-0123548.

PR 23-MAR-1999: 99US-0125788.

PR 23-MAR-1999: 99US-0126264.

PR 29-MAR-1999: 99US-0126785.

PR 01-APR-1999: 99US-0127462.

PR 06-APR-1999: 99US-0128234.

PR 08-APR-1999: 99US-0128714.

PR 16-APR-1999: 99US-0129845.

PR 19-APR-1999: 99US-0130077.

PR 21-APR-1999: 99US-0130449.

PR 23-APR-1999: 99US-0130891.

PR 28-APR-1999: 99US-0131449.

PR 30-APR-1999: 99US-0132048.

PR 04-MAY-1999: 99US-0132407.

PR 05-MAY-1999: 99US-0132485.

PR 06-MAY-1999: 99US-0132486.

```
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134371.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 17-JUN-1999: 99US-0139453.
PR 18-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 21-JUN-1999: 99US-0139763.
PR 22-JUN-1999: 99US-0139763.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141284.
PR 01-JUL-1999: 99US-0141842.
PR 02-JUL-1999: 99US-0142154.
PR 06-JUL-1999: 99US-0142205.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 16-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
```

```

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155458.
PR 29-SEP-1999; 99US-0155456.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

```

```

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160740.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 6.7; Score 32.8; DB 21; Length 533;
Best Local Similarity 53.7; Pred. No. 1.9;
Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```

```

Qy 348 tgaccatctcctgacatctctgctggttctactgctcttggcttctgctggttgac 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 tgatctcatatgaacatctgctgctgctgctgctgctgctgctgctgctgctg 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 408 ggagagagagagagagagagagagagagagagagagagagagagagagagagagag 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 ggagagagagagagagagagagagagagagagagagagagagagagagagagag 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 tgcctcatgaaga 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 cgatlaagtgca 375

```

## RESULT 11

AAV19497 standard; DNA: 585 bp.

AAV19497:

18-SEP-1998 (first entry)

Cytochrome P45011d6 encoding DNA 847-1431.

Human; cytochrome P45011d6; liver/kidney microsomal; LKM-1; antibody;

autoimmunohepatitis type II; A1H type II; hepatitis C virus; ss.

Hepatitis C virus.

Key 1.585; function/analogs

FT CDS /transl\_except (pos:316..318,aa:xa)

FT /note "Xaa is unspecified"

PN JP10101698-A.

PD 21-APR-1998.

PF 24-SEP-1996; 96JP-0273055.

PR 24-SEP-1996; 96JP-0273055.

PA (HOKE-) HOKEN KAGAKU KENKYUSHO KK.

DR WPI: 1998-292131/26.

DR P-PSDB; AAW44870.

PT Amino acid sequence for liver/kidney microsomal (LKM) 1 antibody - to

human auto-immuno-hepatitis (A1H) type II and hepatitis C

PS Claim 3; Fig 3; 9pp; Japanese.

CC The present sequence encodes a protein fragment of cytochrome P45011d6









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OR nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:12:49 ; Search time 226.99 seconds  
(without alignments)  
499,870 Million cell updates/sec

Title: US-08-956-991-9\_Copy\_500\_1000

Perfect score: 501  
Sequence: 1 accttcgcgcagcagcagata.....aacacacagacacagatac 501

Scoring table:  
IDENTITY: 100%  
Gap: 10.0, Gapext: 1.0

Search: 351203 seqs, 11323899 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCPUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/backfilltest.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.8	6.7	1171	1	US-08-336-257A-1
2	33.8	6.7	1171	6	5386025-1
3	32.0	6.5	1396	1	US-08-123-161A-11
4	32.6	6.5	1396	1	US-08-484-278-11
5	32.6	6.5	1436	4	US-08-910-864-11
6	31.8	6.3	533	6	5482709-5
7	31.8	6.3	533	6	5482709-5
8	31.2	6.2	1569	2	US-08-145-658D-23
9	30.8	6.1	1050	1	US-08-599-252-81
10	30.8	6.1	1050	1	US-08-436-074-54
11	30.8	6.1	1050	5	PCT-US96-06352-81
12	30.8	6.1	1050	5	PCT-US96-06583-81
13	30.6	6.1	1501	2	US-08-784-651-3
14	30.4	6.1	1566	2	US-08-145-658D-24
15	30.4	6.1	1566	2	US-08-145-658D-13
16	30.4	6.1	1701	1	US-07-863-169A-2
17	30.4	6.1	1701	2	US-08-429-964-2
18	30.4	6.1	1701	2	US-07-935-087-2
19	30.4	6.1	1701	5	PCT-US93-08062-20
20	30.2	6.0	1568	2	US-08-145-658D-20
21	30	6.0	3366	2	US-08-469-802H-1
22	30	6.0	3366	2	US-08-267-803A-1
23	30	6.0	10660	2	US-08-267-803B-8
24	30	6.0	10660	4	US-09-041-886-16
25	29.8	5.9	143	6	5508199-7
26	29.8	5.9	143	2	US-08-728-123A-1
27	29.8	5.9	32207	2	US-08-770-379-20

28	29.8	5.9	32207	4	US-08-757-669A-20	Sequence 20, Appl
29	29.4	5.9	1566	2	US-08-145-658D-22	Sequence 22, Appl
30	29.4	5.9	1571	2	US-08-145-658D-21	Sequence 21, Appl
31	29.4	5.9	252	4	US-09-115-446-1	Sequence 1, Appl
32	29.4	5.9	2526	4	US-09-115-446-5	Sequence 5, Appl
33	29.4	5.9	5822	3	US-08-899-595-4	Sequence 4, Appl
34	29.4	5.9	5822	3	US-08-899-595-5	Sequence 5, Appl
35	29.2	5.8	435	3	US-08-301-162-9	Sequence 9, Appl
36	29.2	5.8	623	4	US-09-043-303-5	Sequence 5, Appl
37	29.2	5.8	1140	1	US-08-424-268-7	Sequence 7, Appl
38	29.2	5.8	1140	5	PCT-US93-10442-7	Sequence 7, Appl
39	29.2	5.8	1664	1	US-07-863-169A-6	Sequence 6, Appl
40	29.2	5.8	1664	2	US-08-429-964-6	Sequence 6, Appl
41	29.2	5.8	1664	4	US-07-935-087-6	Sequence 6, Appl
42	29.2	5.8	1664	5	PCT-US93-08062-6	Sequence 6, Appl
43	29.2	5.8	2546	1	US-08-424-268-19	Sequence 19, Appl
44	29.2	5.8	2546	5	PCT-US93-10442-19	Sequence 19, Appl
45	29.2	5.8	291	1	US-08-188-277B-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-08-336-257A-1  
Sequence 1, Application US/08336257A  
Patent No. 5726015  
GENERAL INFORMATION:  
APPLICANT: Jay, Scott D.  
APPLICANT: Ellis, Steven H.  
APPLICANT: Harpold, Michael M.  
APPLICANT: Campbell, Kevin P.  
TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,257A  
FILING DATE: 07-Nov-1994  
CLASSIFICATION: 433  
ATTORNEY/AGENT INFORMATION:  
NAME: Seddman, Stephanie L.  
REGISTRATION NUMBER: 35,779  
REFERENCE/BOOKET NUMBER: 54898  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0062  
TELEFAX: (619) 238-0062  
INFORMATION FOR SPO ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1171 base pairs  
TYPE: nucleic acid  
STRANDNESS: not relevant  
TOPOLGY: not relevant  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 49..717  
OTHER INFORMATION: /product- "Gamma subunit of animal  
OTHER INFORMATION: calcium channel"  
US-08-336-257A-1  
Query Match 6.7%: score 33.8; DB 1; length 1171;







RESULT 10  
US-08-436-074-54  
Sequence 54, Application US/08436074  
Patent No. 575438  
GENERAL INFORMATION:  
APPLICANT: DRAVNA, DENNIS T.  
APPLICANT: FEDER, JOHN N.  
APPLICANT: GRIKKE, ANDREAS  
APPLICANT: KIMMEL, BRUCE E.  
APPLICANT: THOMAS, WINSTON J.  
APPLICANT: WOLFF, ROGER K.  
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY  
TITLE OF INVENTION: HEMOCHROMATOSIS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08-436,074  
FILING DATE: 08-MAY-1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 9053-0001.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1050 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-436-074-54

Query Match 6.1%; Score 30.8; DB 1; Length 1050;  
St. Local Similarity 53.3%; Pred. No. Z;  
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 334 gctcaagatcctcgtacacatcctcgtacatccctggtcagatctcctggtccttctggtcct 393  
DB 301 gcccaagatatttgatgacgttcgcagctgcctgctgcgtgcgcctgctgcctgcctgcct 360  
QY 394 tctgctggttctgctgag 453  
DB 361 tctgctggttctgctgag 420  
QY 454 ga 455  
DB 421 GA 422

RESULT 11  
PCT-US96-06352-81  
Sequence 81, Application PCT/US9606352  
GENERAL INFORMATION:  
APPLICANT: DRAVNA, DENNIS T.  
APPLICANT: FEDER, JOHN N.  
APPLICANT: GRIKKE, ANDREAS  
APPLICANT: KIMMEL, BRUCE E.  
APPLICANT: THOMAS, WINSTON J.

APPLICANT: WOLFF, ROGER K.  
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY  
TITLE OF INVENTION: HEMOCHROMATOSIS  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06352  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/599,252  
FILING DATE: 09-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 9053-0001.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 81:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1050 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US96-06352-81

Query Match 6.1%; Score 30.8; DB 5; Length 1050;  
Best Local Similarity 53.3%; Pred. No. Z;  
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 334 gctcaagatcctcgtacacatcctcgtacatccctggtcagatctcctggtccttctggtcct 393  
DB 301 gcccaagatatttgatgacgttcgcagctgcctgctgcgtgcgcctgctgcctgcctgcct 360  
QY 394 tctgctggttctgctgag 453  
DB 361 tctgctggttctgctgag 420  
QY 454 ga 455  
DB 421 GA 422

RESULT 12  
PCT-US96-06352-81  
Sequence 81, Application PCT/US9606352  
GENERAL INFORMATION:  
APPLICANT: DRAVNA, DENNIS T.  
APPLICANT: FEDER, JOHN N.  
APPLICANT: GRIKKE, ANDREAS  
APPLICANT: KIMMEL, BRUCE E.  
APPLICANT: THOMAS, WINSTON J.  
APPLICANT: WOLFF, ROGER K.  
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY  
TITLE OF INVENTION: HEMOCHROMATOSIS  
NUMBER OF SEQUENCES: 124  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500





Search completed: March 4, 2002, 05:13:00  
Job time: 17034 sec